3.0 Risk Assessment Framework

The HWC risk analysis completed for the final rule characterizes both human health and ecological risk for the universe of HWC facilities located within the continental United States for the following combustor categories:

- Cement kilns
- Lightweight aggregate kilns
- Commercial incinerators
- On-site incinerators (large and small)
- Waste heat boilers
- Area sources.

Section 3.1 discusses the key components of the analytical approach used for this risk assessment and Section 3.2 describes the modeling process used.

The analytical approach described in this section differs in important ways from the approach used for the risk analysis for the proposed rule. Specifically, there are seven major differences:

- For the final rule risk analysis, 76 facilities were modeled, which is a substantial increase over the 11 facilities modeled for the proposed rule risk analysis. Moreover, the facilities modeled were selected in a statistically meaningful manner so that inferences could be made about the universe of facilities. That is, the 76 facilities modeled are representative of the larger universe.

- For the final rule risk analysis, all human receptor populations were enumerated except for the subsistence scenarios. The human populations for a given receptor were further divided into four age groups to allow risk characterization for children.

- The proposed rule risk analysis located specific residences and farms in the proximity of the modeled facility. In the risk analysis for the final rule, risk to the entire population was evaluated. Results of the modeling are presented as a distribution of exposure and of risk weighted by the affected populations.

- The basic risk results are based on central tendency values for all exposure parameters. The resulting distribution of risk captures most but not all of the variability in exposure and risk. Therefore, the risk analysis for the final rule also contains an assessment of the variability in selected exposure parameters and models their influence on exposure and risk values. This exposure variability assessment was conducted for the important risk-driving exposure pathways.

- The final rule risk analysis includes a multipathway risk analysis for three species of mercury.

- The final rule risk analysis includes a lead analysis. Blood lead levels were modeled for the 0- to 5-yr-old age group. This allowed lead risk levels to be characterized in terms of the number of individuals in the 0- to 5-yr-old age group who exceeded a blood lead level of concern.

- Finally, the risk assessment for the final rule includes a comprehensive screening-level analysis of ecotoxicological risks.

3.1 Analytical Overview

This section provides an overview of the analytical approach used to evaluate both human health and ecological risk for the final rule. Emphasis is placed on introducing those techniques and approaches related to exposure assessment and risk characterization that were developed specifically for the HWC risk analysis.

3.1.1 Facility Selection

A critical requirement in developing the HWC risk analysis methodology was that it allow clear statistical statements to be made concerning the representativeness of the risk results for the universe of HWC facilities (those within the continental United States). The methodology developed for this analysis specifically addressed this representativeness goal by incorporating a facility-specific modeling approach and using stratified random sampling to select the facilities to be modeled.

3.1.1.1 <u>Facility-Specific Modeling Approach</u>. The facility-specific modeling approach combined the site-specific analyses of facility emissions, fate and transport, and exposed receptor populations with national data on exposure factors to generate estimates of exposure and risk.

3.1.1.2 <u>Stratified Random Sampling Approach</u>. The stratified random sampling approach was used to select specific facilities from the HWC universe, which forms the basis of the risk analysis. The HWC universe was stratified according to the combustor categories of interest (e.g., cement kilns and waste heat boilers), and facilities to be modeled were randomly sampled from those strata. The use of random sampling allowed clear statistical statements to be made concerning the representativeness of risk results generated for the modeled facilities (i.e., how representative those results are of the universe of HWC facilities). Sampling error, which results from not having sampled all of the facilities in the universe, could be quantified by placing confidence intervals (reflecting sampling error) around specific risk estimates.

Stratified random sampling was conducted separately for each combustor category and was continued within each category until a sufficient number of facilities had been sampled to provide a 90 percent probability that at least one selected facility was a high-risk facility. With random sampling, a quantitative statistical criterion (i.e., a 90 percent probability of selecting a high-risk facility) could be identified and reflected directly in the selection of facilities.

3.1.2 Exposure Assessment

The exposure assessment examined the exposure of human receptor populations to those constituents released to the atmosphere by HWC facilities that can be quantified. Constituents assessed were

- **Seven congeners of chlorinated dioxin**
  2,4,7,8 - Tetrachlorodibenzo(*p*)dioxin
  1,2,3,7,8- Pentachlorodibenzo(*p*)dioxin
  1,2,3,7,8,9 - Hexachlorodibenzo(*p*)dioxin
  1,2,3,4,7,8, - Hexachlorodibenzo(*p*)dioxin
  1,2,3,6,7,8 - Hexachlorodibenzo(*p*)dioxin
  1,2,3,4,6,7,8 - Heptachlorodibenzo(*p*)dioxin
  1,2,3,4,5,7,8,9 - Octachlorodibenzo(*p*)dioxin

- **Ten congeners of chlorinated furan**
  2,3,7,8 - Tetrachlorodibenzo(*p*)furan
  1,2,3,7,8- Pentachlorodibenzo(*p*)furan
  2,3,4,7,8- Pentachlorodibenzo(*p*)furan
  1,2,3,6,7,8- Hexachlorodibenzo(*p*)furan
  2,3,4,6,7,8- Hexachlorodibenzo(*p*)furan
  1,2,3,4,7,8- Hexachlorodibenzo(*p*)furan
  1,2,3,7,8,9- Hexachlorodibenzo(*p*)furan
  1,2,3,4,6,7,8- Heptachlorodibenzo(*p*)furan
  1,2,3,4,7,8,9- Heptachlorodibenzo(*p*)furan
  1,2,3,4,6,7,8,9- Octachlorodibenzo(*p*)furan

- **Three species of mercury**
  Elemental mercury
  Divalent mercury
  Methylmercury

- **Eleven metals that were modeled for the proposed rule**

Antimony	Beryllium
Chromium III, VI	Selenium
Arsenic	Cadmium
Lead	Silver
Barium	Thallium
Nickel	

- **Three additional metals modeled for the final rule**
  Cobalt
  Copper
  Manganese

- **Particulate matter**
  PM_{10}
  $PM_{2.5}$

- **Hydrochloric acid**

- **Chlorine gas**

The HWC risk analysis assessed human health risks for various receptor populations. A critical component of the analysis was the location and density of receptor populations relative to the modeled facilities. Air modeling results for a given facility define a pattern of air concentration and deposition values for constituents of concern within the study area. For this final rule analysis, these detailed air model results were linked to spatially refined population estimates and land use characteristics. Specifically, each modeled study area (comprising the modeled facility and the surrounding 20-km radius area) was divided into 16 sectors using four concentric rings combined with a north-south and east-west transect (see Section 4.3).

A geographic information system (GIS) platform was used to enhance 16-sector spatial resolution since key site attributes linked to exposure could be defined at the sector level. These attributes were: air model results, density of receptor populations, topography, waterbodies, watersheds, soils, and land use type. The ability to define these attributes at the sector level provided the level of resolution required to generate sector-level projections of both individual and population risk for the human health component of the analysis as well as sector-level characterization of potential ecological impacts.

To further enhance exposure assessment with regard to human health for the final rule, four separate age groups were used to characterize risk. The use of four age groups (0-5, 6-11, 12-19, and >19 years) allowed age-dependent differences in exposure parameters to be reflected in both exposure assessment and risk characterization. The U.S. Census contains data with sufficient age-group resolution to allow the generation of population estimates at the sector level for these age groups. Also included in the analysis for selected constituents (e.g., dioxins and furans) is an assessment of nursing infants exposed via maternal milk.

3.1.3 Human Health Risk Characterization

The risk assessment methodology implemented for the final rule characterized risks to both human and ecological receptors located within 20 km of facilities within the HWC universe. There was no consideration of risks resulting from atmospheric constituents transported beyond the 20-km study areas. Inferences about risks posed by the universe of HWC facilities were made based on risk estimates generated for the subset of modeled facilities. The statistical analysis that applied facility sample weights and population weights to the sector-level risk results based on a stratified random sample of facilities was conducted using SUDAAN, a statistical analysis software package developed by RTI. All risk estimates generated for the final rule are presented according to the key combustor categories.

Because risks were generated at the sector level through the use of the 16-sector template, sector-level risk estimates form the basis for projecting both individual and population risk estimates for the human receptor as well as ecological risk estimates. The HWC analysis was designed to characterize two broad categories of human health risk: individual and population. For individual risk, emphasis was placed on characterizing distribution of individual risk within the receptor population (e.g., risk to the 50^{th} percentile individual within the population and risk for the 90^{th}, 95^{th}, and 99^{th} percentile individual). Population risk was evaluated both for local populations (those individuals residing within 20 km of an HWC facility) and the national population (those individuals who consume agricultural commodities produced within 20 km of an HWC facility but who reside outside the 20-km study area).

A significant enhancement in individual risk characterization implemented for the final rule was the use of population-weighted individual risk distributions for the identification of specific individual risk percentiles. For the final rule, population-weighted individual risk estimates were used as the basis for a cumulative individual risk distribution rather than unweighted sector estimates. Each sector-level individual risk estimate was first weighted to reflect the number of individuals from the receptor population of interest located within that sector. This approach allowed the distribution of individuals across a study area to be reflected in the cumulative risk distributions used to identify specific individual risk percentiles.

The population-weighted individual risk approach can be applied only to enumerated receptor populations. For those populations that could not be enumerated using Census data (e.g., subsistence scenarios), unweighted sector-level individual risk estimates were used to form the cumulative risk distributions from which individual risk percentiles were selected.

Individual risk estimates were generated for those constituents with carcinogenic effects using standard risk assessment techniques. For noncancer effects, exposures were compared to a reference dose and expressed as a ratio or hazard quotient. In addition, for lead, individual exposures in children were generated as body burden levels in blood. Furthermore, an incremental margin of exposure was used to assess the potential for noncancer effects for dioxin. This was done for infants exposed to dioxin through breast milk as well as for the full set of receptor populations and age groups considered in this risk analysis.

Individual risk estimates were generated for those constituents identified as having carcinogenic effects based on the lifetime average daily dose combined with a cancer slope factor. The CSF is an upper bound estimate of the probability of an individual developing cancer over a lifetime per unit intake of a contaminant. Overall cancer risk was estimated assuming additivity.

Individual risk estimates were generated for those constituents identified as having non-cancer effects based on the ratio of the average daily dose (ADD) to a reference dose or the ratio of annual average air concentrations to a reference concentration. The ratio representing individual risk estimates is the hazard quotient. The reference dose is an estimate of the average daily dose that is without appreciable risk of deleterious effects during a lifetime. An overall hazard index was generated as the sum of the constituent-specific hazard quotients.

The HWC risk analysis completed for the final rule characterizes population risk resulting from human exposure to constituents deposited within HWC study areas. The selection of population risk categories for the final rule focused on those health effects that could be quantified. With regard to carcinogenic risk, two types of statistical cancer incidence estimates are presented:

- Agricultural commodity statistical cancer incidence analysis estimates the number of statistical cancer incidence cases occurring nationally as a result of the public's consumption of beef, milk, and pork raised within HWC study areas. These agriculture commodities have been impacted by dioxin released from their local HWC facility.

- Local statistical cancer incidence analysis estimates the number of statistical cancer cases occurring strictly within the HWC study areas as a result of local (i.e., individuals living within study areas) exposure to all modeled carcinogens. This analysis considers all modeled exposure pathways including the ingestion of home-produced agricultural commodities.

Besides these cancer population risk analyses, the HWC risk analysis also included population risk analyses, including the number of children exposed to lead above health-based levels and adverse health effects resulting from inhalation of PM_{10} and $PM_{2.5}$.

In addition to the above quantitative population risk categories, semiquantitative population risk statements are also provided for exposure of recreational fishers to mercury through fish ingestion. This population risk category estimates the number of recreational fishers potentially engaging in fishing activity in at-risk waterbodies (i.e., modeled waterbodies with individual risk levels for fish ingestion above the health benchmark level [HBL] for methylmercury).

3.1.4 Ecological Risk Characterization

The ecological risk component of the HWC analysis assessed the potential for adverse impacts to both aquatic and terrestrial receptors as a result of exposure to modeled constituents released from HWC facilities. The ecological risk analysis considered impacts only to ecological receptors located primarily within study areas. This analysis was based on the development of criteria (e.g., protective media concentrations) that, in turn, were based on ecological benchmarks (e.g., no observed adverse effects levels or NOAELs). Modeled media concentrations (including soil, surface water, and sediment) were compared to these ecological criteria at the sector level to determine whether the potential for ecological impacts existed within a given study area (i.e., do HQs exceed unity).

For dioxin, a different approach was taken to address ecological risks in aquatic systems. Instead of comparing modeled water concentrations to media-specific ecotoxicological criteria, the dietary intake of dioxins (expressed as toxicity equivalents or TEQs) for receptor organisms was compared directly to the ecotoxicological benchmarks for 2,3,7,8,-TCDD. This approach allowed the assessment of ecological exposures for all 2,3,7,8-chlorine-substituted congeners, taking into consideration the differential toxicity and bioaccumulation of different congeners in the aquatic food chain.

A critical factor in determining the significance of HQ exceedances is the spatial pattern of those exceedances. The use of the 16-sector template allowed spatial patterns to be identified and evaluated for their potential ecological significance.

Although this ecological analysis was based on a comprehensive set of ecological criteria, it is a screening-level analysis designed to identify the **potential** for adverse impacts to ecological receptors and does not provide quantitative results as does the human health evaluation.

As with the human health analysis, ecological risk results generated for modeled HWC facilities are facility-sample-weighted to represent the universe of HWC facilities (see discussion in Section 3.1.3).

3.2 Overview of Modeling Process

The modeling process used in this human health and ecological risk assessment of HWC facilities involves a series of steps beginning with selection of HWC facilities to be modeled and ending with characterization of human and ecological risks. The purpose of this section is twofold: (1) to provide an overview of the steps involved in the modeling process and (2) to provide a map to the discussion of modeling methodologies presented in subsequent sections of this report.

Figure 3-1 shows the steps involved in the modeling process used and groups those steps into six broad categories:

- Characterizing modeled facilities
- Determining environmental media concentrations
- Determining food chain concentrations
- Calculating human intake and dose
- Characterizing human health risks
- Characterizing ecological risks.

These six categories define the main components of the modeling process. Figure 3-1 also cross references each of these components to the appropriate section of this document containing greater detail.

3.2.1 Characterizing Modeled Facilities

The HWC risk assessment methodology is based on a facility-specific modeling approach; therefore, the first step in the modeling process is to define the universe of all HWC facilities and then select the facilities to be modeled from this universe. Stratified random sampling was used to select facilities for the final rule, which resulted in 66 facilities being selected. These 66 were combined with 10 of the 11 facilities modeled for the proposed rule, resulting in 76 facilities modeled for this risk analysis. These 76 facilities represent the universe of incinerator, cement kiln, and lightweight aggregate kiln source categories (see Section 4.1.1).

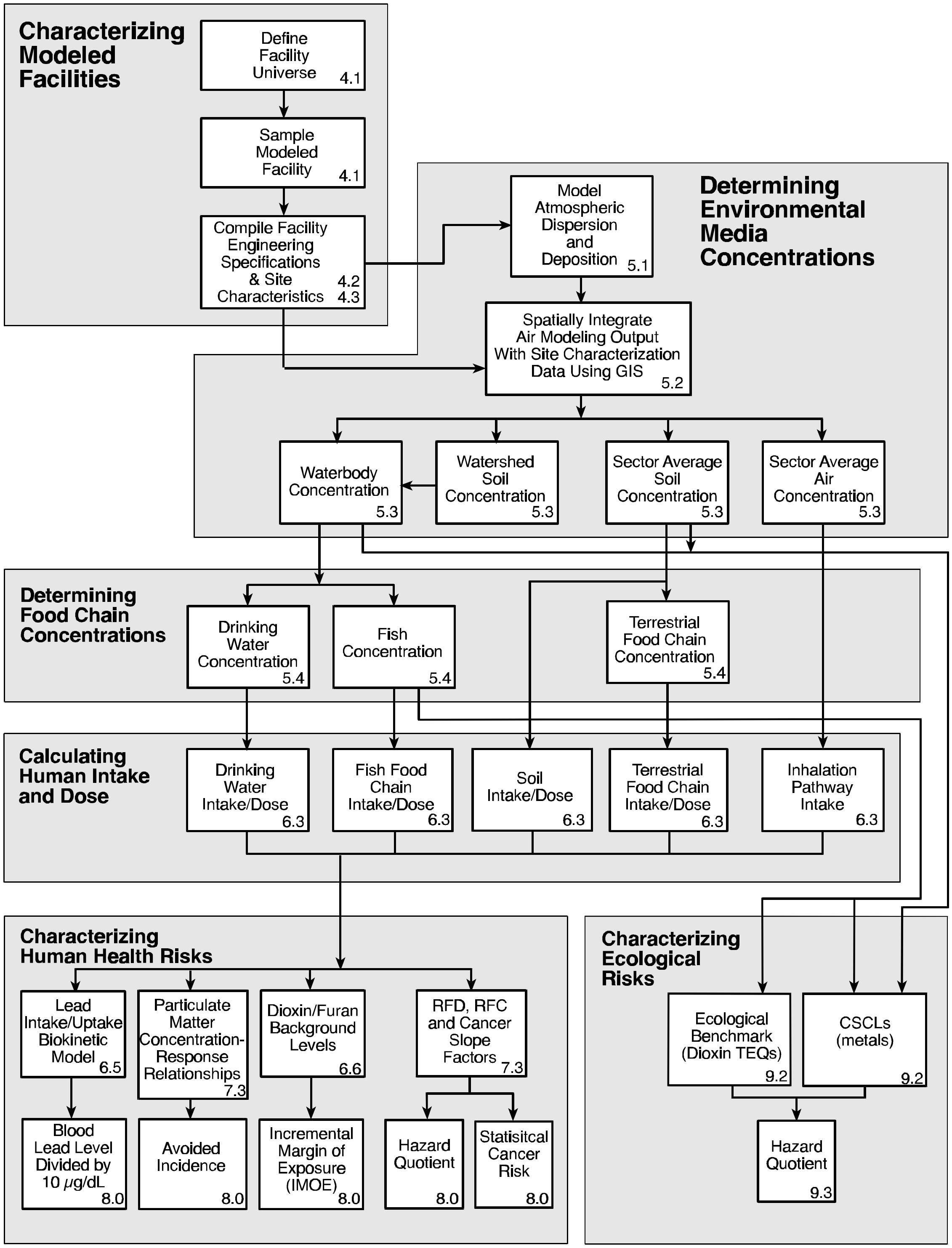

Figure 3-1. Overview of risk assessment framework.

3.2.2 Determining Environmental Media Concentrations

Air dispersion and deposition modeling was conducted using EPA's Industrial Source Complex Model - Short Term Version 3 (ISCST3) to arrive at normalized air concentrations and deposition fluxes (see Section 5.1). Modeling was based on a 1-g/s emission rate (a normalized emission rate). The air modeling grid data were then converted using a GIS into average normalized values for geographic features in the study area: sectors, watersheds, and waterbodies (Section 5.2). These normalized values were then combined with facility-specific emissions data to calculate waterbody concentrations, watershed soil concentrations, sector air concentrations, and sector soil concentrations (Section 5.3). Sector soils, watershed soils, and waterbody concentrations were modeled using the 1993 Addendum to the Methodology for Assessing Health Risks Associated with Indirect Exposure to Combustor Emissions for all constituents except mercury (U.S. EPA, 1993). Mercury species in soils and waterbodies were modeled in two different ways. Mercury modeling for the aquatic food chain pathway (watershed-waterbody-fish tissue) was done using IEM-2M based on the 1997 *Mercury Study Report to Congress* (U.S. EPA, 1997) (see Section 5.3.3.2). The IEM-2M methodology was not used to model mercury in sector soils and the drinking water pathway; they were modeled using a version of the IEM-2 methodology that was modified specifically for this risk assessment (see Appendix F).

3.2.3 Determining Food Chain Concentrations

The media concentrations obtained in the previous step were used to calculate food chain concentrations as follows (Section 5.4):

- Terrestrial food chain concentrations were based on air and soil concentrations for each sector.

- Drinking water concentrations were based on waterbody concentrations. The majority of modeled facilities had at least one waterbody identified as the drinking water source for a community.

- Fish tissue concentrations were based on modeled waterbody concentrations for recreational and subsistence fishers and on farm pond concentrations for subsistence farmer populations.

Media and food chain concentrations calculated in the previous step were combined with intake rates, which were generated for each of the modeled pathways to produce constituent-specific exposure estimates for those pathways. Intake rates refer to the modeled rates of ingestion or inhalation that were generated for specific types of media or food commodities (e.g., incidental ingestion rates for soil generated for the adult commercial beef farmer). Exposure estimates, which were calculated separately for each constituent/pathway combination, represent the rate of exposure to a specific constituent that results from the ingestion or inhalation of a specific type of media or food commodity.

3.2.4 Modeling Human Exposure

The HWC risk analysis assessed exposure for a number of receptors, each of which was modeled using a suite of exposure pathways designed to capture the receptor's activity/behavior pattern. Receptors modeled in the analysis and their pathways are listed in Table 3-1. Receptors are defined as follows:

- Residents: individuals residing within HWC study areas
- Home gardeners: individuals residing within HWC study areas who engage in home gardening activity
- Recreational fishers: individuals residing within HWC study areas who engage in recreational fishing activity
- Commercial beef farmers: individuals who operate commercial beef farms within HWC study areas
- Commercial pork farmers: individuals who operate commercial hog farms within HWC study areas
- Commercial dairy farmers: individuals who operate commercial dairy farms within HWC study areas
- Commercial produce farmers: individuals who operate commercial produce farms within HWC study areas
- Subsistence fishers: individuals who reside within HWC study areas and obtain all of their dietary fish intake from home-caught fish
- Subsistence farmers: individuals who reside within HWC study areas and obtain all of their dietary intake from home-produced food items.

To gain greater resolution in assessing exposure for the receptors listed above, each receptor was further differentiated into four age groups (i.e., 0-5, 6-11, 12-19, and >19 yr), and separate exposure estimates were generated for each age group. In addition, for dioxins, exposure to human infants from maternal milk was modeled.

Exposure was calculated based on intake values for each of the pathways presented in Table 3-1. Two different types of exposure estimates were generated depending on the type of health effect being characterized. Carcinogenic health effects are characterized using exposure estimates that are averaged over the lifetime of the individual (LADDs). Noncancer effects are characterized using exposure estimates that are averaged over the relevant averaging period (nominally 1 year) during which the exposure occurs (ADDs). All exposure estimates are expressed as daily doses for a specific constituent normalized for the body weight of the receptor (i.e., mg *constituent*/kg *body weight* per day or mg/kg-d).

Table 3-1. Receptors Modeled by Pathways

Receptors	Inhalation of ambient air	Incidental soil ingestion	Ingestion of drinking water	Ingestion of home-produced fruits and vegetables	Ingestion of home-caught fish	Ingestion of home-produced beef	Ingestion of home-produced pork	Ingestion of home-produced milk	Ingestion of home-produced chicken
Residents	✓	✓	✓						
Home gardeners	✓	✓	✓	✓					
Recreational fishers	✓	✓	✓		✓				
Commercial beef farmers	✓	✓	✓			✓			
Commercial hog farmers	✓	✓	✓				✓		
Commercial dairy farmers	✓	✓	✓					✓	
Commercial produce farmers	✓	✓	✓	✓					
Subsistence fishers	✓	✓	✓		✓				
Subsistence farmers	✓	✓	✓	✓	✓	✓	✓	✓	✓

3.2.5 Characterizing Human Health Risks

The HWC risk analysis assessed risks for a number of different human health effects, including cancer, noncancer effects, health effects from lead, health effects from PM, and noncancer effects associated with dioxin/furan exposure. A combination of both individual and population-level risk descriptors were used in characterizing risks for these health effects.

3.2.5.1 <u>Cancer</u>. Individual cancer risk was evaluated by multiplying the LADD estimates generated for each receptor/pathway by the appropriate cancer slope factor. Cancer slope factors were derived from either human or animal data and relate the level of exposure to a particular constituent to the lifetime excess cancer risk that results from that exposure. In developing cancer slope factors, the relationship between exposure and risk is generally assumed to be linear with the slope factor representing the upper bound on the slope of the dose-response curve in the low-dose region where modeled human exposure typically occurs. Total individual cancer risk was determined for each receptor, assuming additivity across constituents.

In the HWC risk analysis, population-level cancer risk is characterized using annual lifetime cancer incidence estimates. These estimates represent the excess number of cancer cases predicted to occur due to emissions released from the facility under evaluation during a single model year. Accordingly, annual incidence is estimated by dividing the total lifetime cancer incidence by the exposure duration.

3.2.5.2 <u>Noncancer Effects</u>. Individual noncancer risk for ingestion pathways was evaluated by dividing the ADD estimates generated for each receptor/pathway by the appropriate RfD to produce a hazard quotient. Inhalation pathways were evaluated for noncancer effects by dividing modeled ambient air concentrations for specific constituents by the corresponding RfC to produce inhalation hazard quotients. RfDs and RfCs, both of which can be based either on human or animal data, represent estimates of daily exposure to the human population, including sensitive subgroups, that are likely to be without an appreciable risk of deleterious effects during a lifetime. Ingestion and inhalation hazard indices were generated for each receptor by adding constituent-specific hazard quotients by route of exposure.

3.2.5.3 <u>Health Effects from Lead</u>. Risk resulting from exposure to lead was assessed for the child age group (i.e., 0 to 5 years old) of every receptor population evaluated in the analysis. Risk for this age group was assessed by modeling body burdens (as blood lead levels) and comparing these levels to the level at which efforts aimed at prevention are indicated (i.e., 10 μg lead/dL blood). In addition to characterizing individual risk levels for lead exposure in the modeled receptor populations, this analysis included population risk estimates expressed as the annual excess incidence of elevated blood lead (i.e., above 10 µg/dL).

3.2.5.4 <u>Health Effects from PM</u>. Risk associated with inhalation exposure to particulate matter was evaluated in the elderly and the general population through the use of concentration-response functions derived from human epidemiological studies that describe the incidence of mortality and morbidity avoided annually due to an incremental reduction in PM. The PM analysis generates only population-level risk estimates.

3.2.5.5 <u>Noncancer Effects from Dioxin/Furan Exposure</u>. Potential noncancer risk associated with dioxin/furan exposure is evaluated using an incremental margin of exposure (incremental MOE) approach. With this approach, modeled exposure levels for specific receptors, expressed as 2,3,7,8-TCDD toxicity equivalents (TEQs), were compared to background TEQ exposure levels in the general population and expressed as a ratio. In addition to generating incremental MOE estimates for each of the four age groups within each receptor, this analysis generated incremental MOE estimates for infant receptors who are exposed to dioxin/furan through the ingestion of breast milk. As a measure of hazard, the incremental MOE presumes that background exposures pose only a de minimis level of risk.

3.2.6 Characterizing Ecological Risks

The ecological risk assessment is a screening-level analysis designed to identify the potential for adverse ecological effects. The process is based on current EPA guidelines for ecological risk assessment and begins with the selection of assessment endpoints (i.e., the actual environmental values to be protected).

The assessment endpoints are defined by two key elements: (1) a valued ecological entity such as a wildlife species, and (2) an attribute of that entity that is important to protect (e.g., reproductive fitness). Once the assessment endpoints are defined, ecological receptors that may be susceptible to the chemical constituents released from HWC facilities are selected. These receptors include assemblages of species typical of soil, sediment, and surface water communities as well as representative species of mammals and birds found in most parts of the contiguous United States.

For each constituent, ecotoxicological data were reviewed to derive benchmarks (in units of dose) and ecotoxicological criteria below which adverse ecological effects are presumed to be negligible. Ecological benchmarks derived for representative species of birds and mammals (generally no observed adverse effect levels, or NOAELs) were used to calculate ecotoxicological criteria using the assumption that all food items originate from the same contaminated area. For species associated with aquatic habitats (e.g., riverine), the ecotoxicological criteria are given in units of surface water concentration and include ingestion of contaminated water and biota (e.g., fish and aquatic invertebrates). For species associated with terrestrial habitats, the ecotoxicological criteria are given in units of soil concentration and include ingestion of contaminated soil and terrestrial biota (e.g., vascular plants, earthworms). The ecotoxicological criteria for assemblages of species typical of soil, sediment, and surface water communities were derived using statistical inference on ecotoxicological data on individual species attributed to the community. For all metal constituents evaluated in this analysis, the media-specific ecotoxicological criteria were compared to the media concentrations predicted using the environmental fate and transport model with an HQ approach. The HQ approach is similar to the approach used in noncancer health risk assessment (i.e., HQ > 1 indicates the potential for adverse ecological effects).

For dioxin/furan congeners in aquatic systems, a toxicity equivalency concentration approach was used so that congener-specific differences in toxicity and bioaccumulation could be considered. Consequently, the HQ approach for dioxin compared the predicted TEC dose (as an administered dose) to the ecological benchmarks for the representative species evaluated in this screening analysis. For the terrestrial system, a soil TEQ concentration, which reflects only the application of TEFs, was compared to a soil ecotoxicological criterion for 2,3,7,8-TCDD, similar to the approach taken for metals. This approach does not consider the differential bioaccumulation of different congeners and, as such, is likely to be exceedingly conservative.

3.3 References

U.S. EPA (Environmental Protection Agency). 1993. *Addendum to Methodology for Assessing Health Risks Associated with Indirect Exposure to Combustor Emissions* (External Review Draft). EPA/600/AP-93/003. Exposure Assessment Group, Office of Health and Environmental Assessment, Washington, DC.

U.S. EPA (Environmental Protection Agency). 1997. *Mercury Study Report to Congress. Volume III - Fate and Transport of Mercury in the Environment*. EPA 452/R-97/005. Office of Air Quality Planning and Standards and Office of Research and Development, Washington, DC.

4.0 Characterization of Modeled Facilities

The risk assessment for the final rule is based on a facility-specific modeling approach. This section presents the methodology used to select modeled facilities and obtain the site data required to characterize those facilities. Section 4.1 describes the approach used for selecting modeled facilities including the definition of the HWC facility universe. Section 4.2 describes the facility-specific engineering and annual emissions data used in conducting air modeling for each of the modeled facilities. Section 4.3 presents the methodologies used to obtain site data for the study area surrounding each of the modeled HWC facilities including delineation of key topographical features and estimation of human and livestock populations. Figure 4-1 diagrams the relationships between specific analytical tasks related to facility characterization.

4.1 Selection of Modeled Facilities

This section presents the methodology used to define the HWC facility universe and randomly select facilities modeled for risk analysis.

4.1.1 Facility Universe

A critical step in developing the HWC risk analysis involved defining the facility universe that the risk analysis would represent. This universe was developed initially as part of the proposed rule-making effort for the HWC risk analysis. After the initial HWC facility universe had been defined, it was updated to reflect new information on facility closures and entrants to the market. In addition, in the fall of 1997, site visits were made to state environmental and EPA Regional offices to identify additional information that could be used to update the facility universe (e.g., changes in the operational status of existing facilities or identification of new facilities). The HWC facility universe used for the final rule reflects both the public comments and the information gathered during this data collection effort. It includes all HWC facilities located within the continental United States that were operational in 1997. For a more detailed discussion of the facility universe, the reader is referred to *Assessment of the Potential Costs, Benefits, and Other Impacts of the Hazardous Waste Combustion MACT Standards* (U.S. EPA, 1999a).

Facilities outside the continental United States were not included in the facility universe because critical data used in site characterization (e.g., U.S. Census data, Census of Agriculture data, and GIS land use coverage data) were not readily available for them. Therefore, the risk

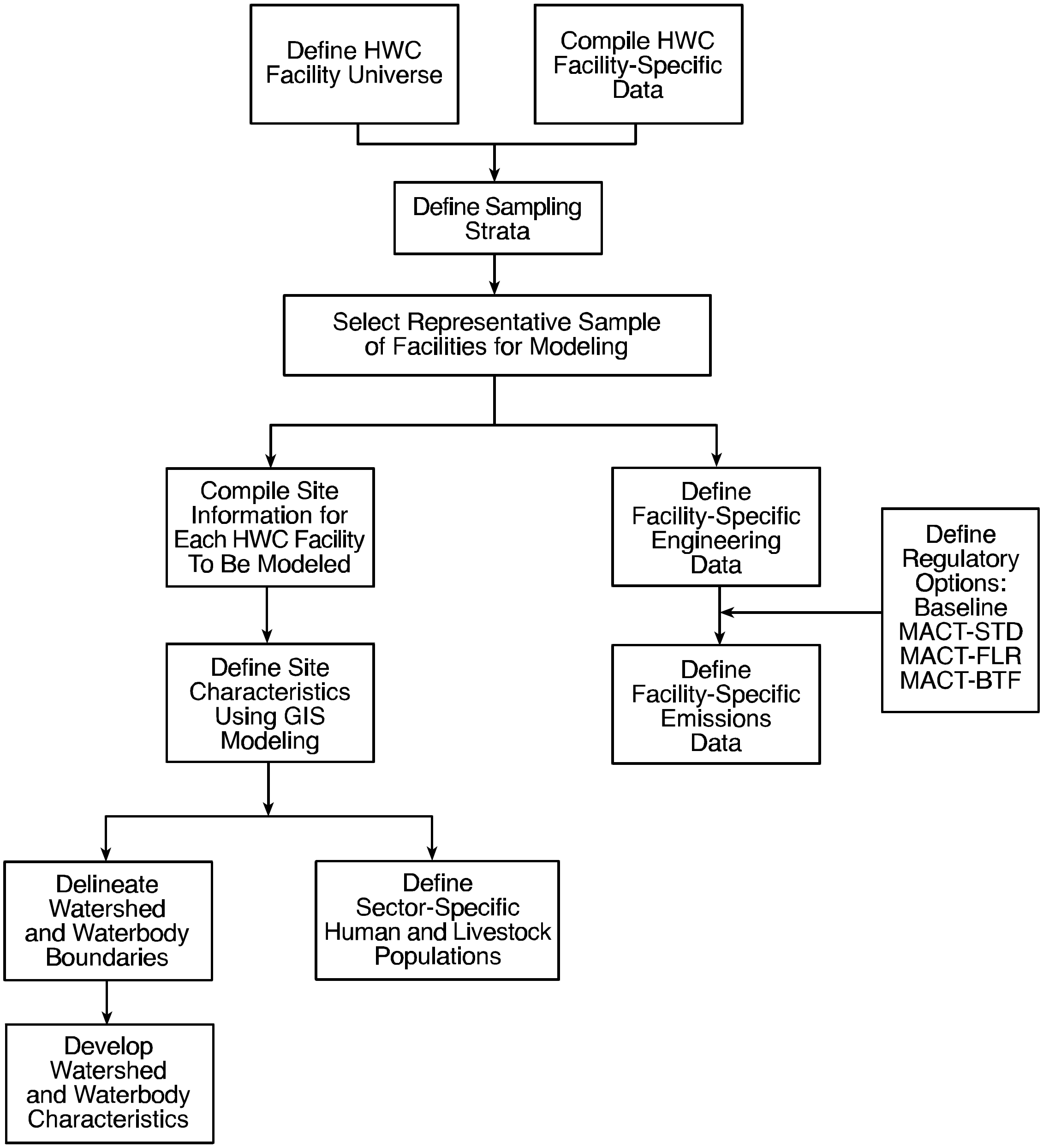

Figure 4-1. Overview of analytical tasks completed for facility characterization.

assessment applies only to those facilities located in the contiguous United States and not to facilities located outside the contiguous United States such as Puerto Rico and Johnson Altoll[1].

4.1.2 Facility Categories

Combustor facilities contained in the HWC universe fall into one of three source categories:

- Cement kilns
- Lightweight aggregate kilns
- Incinerators.

Because the facilities in each of these source categories are linked to specific commercial activities, they tend to share more operational attributes with other facilities in their particular category than with facilities in other categories. Therefore, in evaluating the potential benefits associated with establishing emissions control standards for HWC facilities, EPA initially stratified the facility universe into categories based on these three source categories and separately evaluated the benefits for each. The proposed rule presented risk results for these three source categories.

EPA retained these three source categories as the basis for the final rule analysis. To provide greater resolution in identifying those facility attributes that are correlated with specific categories of risk, however, EPA further stratified the HWC universe by adding several combustor categories to the analysis for the final rule. Some of these new combustor categories are mutually exclusive (e.g., on-site and commercial incinerators), while others extend across several different categories to group facilities that share a particular operational attribute (e.g., waste heat recovery boilers). The following combustor categories have been added for this analysis:

Combustor Categories added for the Final Rule

- Commercial incinerators
- On-site incinerators
- Waste heat boilers
- Area sources

- **Commercial incinerators:** Function specifically as commercial facilities that earn revenue by burning hazardous waste. As such, the incinerators in this combustor category are often larger (i.e., higher throughput) and burn a greater variety of wastes than those in the on-site category. General differences between the commercial and on-site incinerators with regard to facility attributes (e.g., emissions rates and stack parameters) raised interest in stratifying the incinerator category to determine whether the different incinerator categories could be linked to specific patterns of risk.

[1] One small on-site incinerator facility located in Alaska (AK0000094888) was included in the facility universe, despite the fact that it is not located within the contiguous United States. Inclusion of this facility in the sample frame for small on-site incineration does not introduce significant error due to size of the small on-site incineration facility universe.

- # **On-site incinerators:** As part of a larger commercial manufacturing operation, handle hazardous wastes generated specifically by that operation (these facilities do not burn wastes from other companies for profit). Because on-site facilities play a support role and are not dependent on earning profit through hazardous waste combustion, they are often smaller than commercial facilities (their size is dependent on the type of operation they support) and burn a limited variety of wastes. To gain additional resolution in identifying facility attributes linked to risk, the on-site incinerator combustor category was further stratified for the final rule into large on-site incinerators (those with stack gas exhaust volumes greater than 20,000 acfm) and small on-site incinerators (those with stack gas exhaust volumes less than 20,000 acfm).

- # **Waste heat boilers:** Recover excess heat generated in the incineration process as a thermal source for industrial applications rather than releasing it directly to the environment. Only a subset of incinerators have WHBs—cement kilns and LWAK facilities do not. Concerns have surfaced that the operating parameters associated with waste heat boilers may result in greater dioxin/furan formation. Therefore, the WHB category was selected for inclusion in the final rule. Because dioxin/furan formation is the focus for this combustor category, all those risk results involving dioxin-TEQ have WHBs broken out as a separate category.

- # **Area sources:** Facilities with relatively low emission rates of HAPs (facilities with relatively high HAP emission rates are major sources). The Clean Air Act definition of an area source was used in the HWC risk analysis to identify area sources: those facilities having an emission rate for a single HAP of less than 10 tons per year or an emissions rate for combined HAPs of less than 25 tons per year. The area source stratification was included in the HWC risk analysis because area sources are not always subject to MACT standards. To gain greater resolution in evaluating area sources, these facilities were further stratified for purposes of the HWC risk analysis into area source cement kilns and area source incinerators (no area source LWAKs were identified). Because the statutory definition of an area source is based on total facility (industrial complex) emissions, it was not possible to distinguish on-site incinerators located at small industrial complexes that are classified as area sources from on-site incinerators located at large industrial complexes that are classified as major sources. Therefore, most on-site incinerators were excluded from the area source incinerator category.

4.1.3 Facility Definition

For the purpose of this risk analysis, a facility is defined as an industrial complex consisting of one or more hazardous waste combustion units (e.g., incinerators, cement kilns) vented through one or more stacks. For facilities with more than one combustion unit and more than one stack, each stack was modeled separately[2]. Subsequently, the air concentrations and deposition resulting from the emissions from these combustion units/stacks were summed to provide air concentration and deposition values for the total facility. Exposure and risk values were attributed to this combined facility impact. Therefore, for this analysis, the hazardous waste combustion units are described in terms of facilities, and risk results are reported accordingly.

4.1.4 Facility Sample Size

The proposed rule included risk characterization for a purposive sample of 11 HWC facilities. These 11 facilities were selected to provide coverage for the following factors: (1) HWC combustor categories being considered, (2) location of the HWC facilities (land use, topography, meteorological conditions), and (3) facility attributes (e.g., stack gas exhaust volume). Comments to the proposed rule identified the need for a modeled facility selection strategy that was more representative of the universe of HWC facilities than the original 11 HWC facilities (10 of which were retained in the final rule risk analysis)[3]. This requirement resulted in an approach for the final rule that utilized stratified random sampling for the selection of additional modeled facilities. This approach allowed statistical statements to be made regarding representativeness of the risk analysis.

The sample design chosen for the final rule was a stratified, one-stage cluster sample, for which the facilities were selected without replacement. The facilities were considered clusters since the final sampling units were the 16 sectors within each facility study area (see Section 4.3). The facility sampling strata correspond to the six combustor categories of interest:

- Cement kilns
- Lightweight aggregate kilns
- Commercial incinerators
- Large on-site incinerators
- Small on-site incinerators
- Waste heat boilers (a subset of incinerators).

Area sources were not treated as separate strata for the purpose of sampling due to difficulties in defining area source universe.

Sample sizes for each combustor category were based on the goal of having a 90 percent probability of selecting a facility from the top 10 percent of facilities within a given combustor category with regard to risk (i.e., a 90 percent probability of having included a "high-risk" facility in the sample). Table 4-1 presents the sample sizes established for each combustor category and the resulting probabilities for selecting at least one high-risk facility from that combustor category. Because waste heat boilers are a subset of the incinerators (but were

[2] There are a few cases in which an industrial complex has more than one combustion unit. If these combustion units do not all belong to the same source category, the emissions were apportioned to the different source categories and the industrial complex was treated as separate facilities, one for each of the source categories coexisting at the industrial complex.

[3] The 11th facility is undergoing RCRA closure and is no longer burning hazardous waste.

Table 4-1. Hazardous Waste Combustion Facility Stratum Sizes and Sample Sizes

Combustion Facility Category	Stratum Size	Random Sample Size	Original Sample Size	Total Sample Size	High-End Sampling Probability[a]
Cement Kilns	18	10	5	15	98
Lightweight Aggregate Kilns	5	3	2	5	100
Commercial Incinerators					
Including Waste Heat Boilers	20	11	2	13	97
Excluding Waste Heat Boilers	12	7	2	9	95
Large On-Site Incinerators					
Including Waste Heat Boilers	43	17	1	18	94
Excluding Waste Heat Boilers	36	15	0	15	90
Small On-Site Incinerators					
Including Waste Heat Boilers	79	25	0	25	96
Excluding Waste Heat Boilers	65	16	0	16	88
Incinerators with Waste Heat Boilers					

[a]Probability that a facility that lies in the upper 10% of the distribution of risk will be sampled.

sampled as an independent category), information for incinerators is presented in Table 4-1 for each incinerator category as a whole (with waste heat boilers included), each incinerator category without waste heat boilers included, and waste heat boilers as a whole (aggregated across the three incinerator categories). Sampling was conducted separately to provide coverage for each of these different incinerator/waste heat boiler combinations, and risks were generated as separate results for each of these categories.

Because of difficulties in defining the area source universe, area sources were not specifically targeted for sampling, and no specific sample size was considered. The reason for this is that the statutory definition of major sources versus area sources under Section 112 of the CAA is based on total facility-wide emissions of hazardous air pollutants. Specifically, those industrial complexes emitting greater than 10 tons of any one hazardous air pollutant or greater than 25 tons of multiple hazardous air pollutants per year are considered major sources. To define an area source under this definition, information about the industrial complex in which an on-site incinerator is located is needed. Such information was not readily available, making it impossible to adequately characterize the area source universe and, therefore, to define the sampling frame. Because area sources are of interest, however, inferences were made regarding exposure and risk based on those incinerators that could be identified and had otherwise been sampled[4]. For cement kilns, all area sources had been sampled and, therefore, all were used for making such inferences.

In determining the sample size and allocation, a large enough number of sites from each stratum (combustor category) were selected so that at least one of the sites posing the greatest risk was included in the sample. To define what is meant by “the greatest risk,” some number of sites in each stratum were specified. For example, if the N_h sites in the h-th stratum were to be ordered from lowest to highest risk, then some number $N_h^* < N_h$ of sites at the top of the list could be identified as posing the greatest risk. Given N_h^*, the problem becomes one of determining the smallest stratum-level sample size, n_h, that will provide a specified probability of including at least one of these sites. The probabilities are given by

$$\text{Prob}\left\{N_h^* \geq 1 \in S\right\} = 1 - \frac{\binom{N_h - N_h^*}{n_h}}{\binom{N_h}{n_h}} \tag{4-1}$$

where $\text{Prob}\left\{N_h^* \geq 1 \in S\right\}$ means the probability associated with having at least one high-risk facility, N_h, in the sample. What remains is a numerical exercise to determine the smallest value n_h that will provide the specified probability.

The sample size solutions shown in Table 4-1 are obtained by defining $N_h^* = 0.10\,N_h$ and requiring $\text{Prob}\left\{N_h^* \geq 1 \in S\right\} \geq 0.90$. That is, a large enough stratum-level sample size was required to provide a 90 percent chance of including at least one facility from the top 10 percent of facilities with respect to risk.

4.1.5 Facility Sampling

The 11 modeled facilities from the proposed rule (10 of which were retained for the final rule) had been selected purposively, which complicated their inclusion in the risk characterization for the final rule. From a statistical standpoint, however, these 10 facilities were considered along with facility selection conducted for the final rule. Therefore, the 10 facilities evaluated for the proposed rule were defined as certainty samples (had a 100 percent chance of being selected), and the remaining HWC facilities (minus the 10) were used to construct the sampling frame for the stratified random sample.

The sample of facilities for the final rule were randomly selected within each stratum. During facility sampling, two unanticipated circumstances arose that complicated the sample design and sample selection:

[4] Area source incinerators that could be identified included commercial incinerators and on-site incinerators at U.S. Department of Defense installations.

- # Information obtained from state/EPA Regional offices and reviewed after sample selection had started indicated the need to make changes in facility status (e.g., combustor category classification and operational status).

- # After sample selection had been initiated, the decision was made to include waste heat boilers as an analysis domain.

One development that impacted sample selection (the change in facility classification and operational status) meant that the original sampling frame used for sample selection was not representative. Specifically, the sampling frame had some facility type misclassifications, contained some ineligible facilities, and was missing several eligible facilities that were identified during the review of information obtained from states/EPA Regions. After cleaning the sampling frame and recalculating the coverage probabilities, two more supplemental strata were created to increase the coverage for waste heat boilers and large on-site incinerators to the target goal of having a 90 percent probability of selecting a high-risk facility. The decision to include waste heat boilers as a separate analysis domain resulted in the construction of an additional supplemental stratum, since the number of waste heat boilers selected during initial sample selection (i.e., before waste heat boilers were identified as a separate stratum) did not provide an adequate coverage probability.

Table 4-2 presents the frame sizes and sample sizes by sampling strata. The frame and sample sizes exclude facilities that were later determined ineligible. Strata 1 through 6 are associated with the initial sample of 68 facilities from the total of 159 facilities within the original frame. As referred to earlier, the supplemental sample of two facilities was selected in stratum 7 to increase the sample of waste heat boilers. The frame for stratum 7 included all the facilities classified as waste heat boilers at that time that were not previously selected in strata 1 through 6.

Cleaning up the sampling frame involved

- # Correcting previously misclassified combustor category classifications

- # Removing ineligible facilities

- # Adding six new facilities not listed on the original frame (bringing the universe total to 165).

After the sampling frame was corrected, additional waste heat boilers and large on-site incinerators were sampled to provide sufficient coverage for these combustor categories. Specifically, additional waste heat boilers were sampled from stratum 8, which contained all the waste heat boilers not selected in strata 1 through 7, and additional large on-site incinerators were sampled from stratum 9, which contained all the large on-site incinerators not selected in strata 1 through 8.

Table 4-2. Frame and Sample Sizes

Facility Stratum	Number Facilities in Frame	Facility Sample Size	Actual Waste Heat Boilers
1. Facilities Evaluated for Proposed Rule (certainty sample)	10	10	1
2. Cement Kilns	13	10	0
3. Lightweight Aggregate Kilns	3	3	0
4. Commercial Incinerators	16	11	4
5. Large On-site Incinerators	36	13	2
6. Small On-site Incinerators	81	21	6
Total	159	68	13
7. Additional Waste Heat Boilers, First Time	19	2	0 (classification error)
8. Additional Waste Heat Boilers, Second Time	16	3	3
9. Additional Large On-site Incinerators	30	3	0
Total	--	76	16

Because three of the six new facilities identified through review of the state/EPA Regional information were small on-site incinerators that were not waste heat boilers, they did not have a chance to be selected during original sample selection, resulting in undercoverage for the small on-site incinerator category. As described in the weighting section, facility poststratification adjustment was used to compensate for inefficiencies in the original sampling frame, including such factors as undercoverage due to not having included viable facilities in the original sampling frame.

The supplemental sampling strata complicated the selection probabilities. Although the task to account for these complications was not trivial, the large sampling rates for the replacement sampling ameliorate the variance-inflating effects of the inefficient sampling. (Note: Both the initial and supplemental sample have relatively high sampling rates.) That is, because the large sampling rates yield very small variances, the variance inflation effects from the inefficient sampling are negligible in comparison. For additional discussion on the effect of sample/population size and inefficient sampling on variance, see Appendix A.

Table 4-3 presents the final set of sampled facilities used in the risk assessment for the final rule.

Table 4-3. Sample Facilities, Classification, and Sampling Weights

Site Type	Site IDs	Company Name	Location	Area	WHB	Adjusted Facility Sampling Weight[a]
CINC	331	Ross Incineration Serv	Grafton, OH	x		1.959
CINC[b]	221	Rollins Environmental Services	Deer Park, TX	x		1.347
CINC	324	Allied Corp.	Birmingham, AL		x	1.521
CINC	325	Aptus	Coffeyville, KS	x		1.959
CINC	333, 612	Trade Waste Incineration	Sauget, IL			0.857
CINC[b]	214	Rollins Environmental Services	Baton Rouge, LA	x		1.347
CINC	601	Laidlaw Environmental Services INC	Clive, UT	x	x	2.479
CINC	486, 487	Ensco, Inc	El Dorado, AR			0.857
CINC	359	Atochem	Carrollton, KY	x	x	2.479
CINC	210, 211, 212	LWD, Inc.	Calvert City, KY			0.857
CINC	A15	BDT Inc.	Clarence, NY	x		1.959
CINC	209	Laidlaw Environmental Services	Roebuck, SC		x	1.521
CINC	A18	Chemical Waste Mgmt	Port Arthur, TX			0.857
CK[b]	401, 402	Ash Grove Cement Company	Chanute, KS			1.019
CK[b]	320	Lafarge	Alpena, MI			1.325
CK	321	Medusa Cement Company	Demopolis, AL	x		1.130
CK	403, 404, 228	Ash Grove Cement Company	Foreman, AR			1.325

(continued)

Table 4-3. (continued)

Site Type	Site IDs	Company Name	Location	Area	WHB	Adjusted Facility Sampling Weight[a]
CK[b]	304	Lone Star Industries	Greencastle, IN	x		0.870
CK[b]	207, 208	Keystone Cement Company	Bath, PA			1.019
CK	305, 335	Medusa Cement	Wampum, PA			1.325
CK	318, 473	Texas Industries	Midlothian, TX			1.325
CK	322, 323	Lafarge	Fredonia, KS			1.325
CK	302	Lafarge	Paulding, OH			1.019
CK	202	Heartland Cement	Independence, KS			1.325
CK[b]	205, 206	Holnam, Inc.	Holly Hill, SC			1.325
CK	204	Holnam, Inc.	Clarksville, MO			1.325
CK	203	Holnam, Inc.	Artesia, MS			1.019
CK	200, 201, 680, 681	Giant Cement Company	Harleyville, SC			1.325
LWAK[b]	311, 312, 336	Solite	Cascade, VA			1.000
LWAK	310, 475	Solite	Brooks, KY			1.000
LWAK[b]	307, 479	Thermalkem (Norlite)	Cohoes, NY			1.000
LWAK	225	Solite	Norwood, NC			1.000
LWAK	313, 314	Solite	Arvonia, VA			1.000
OINC-Large	A62	Texaco Chemical Co.	Conroe, TX			2.328

(continued)

Table 4-3. (continued)

Site Type	Site IDs	Company Name	Location	Area	WHB	Adjusted Facility Sampling Weight[a]
OINC-Large	504	Chevron Chemical	Philadelphia, PA			2.328
OINC-Large	464	BP Chemicals	Lima, OH			2.328
OINC-Large	A43	Occidental Chemical Corp	Niagara Falls, NY			3.243
OINC-Large	463	Miles	Kansas City, MO			1.978
OINC-Large	480, 706	Ciba-Geigy	St. Gabriel, LA			1.978
OINC-Large	915	Eastman Kodak	Rochester, NY			2.328
OINC-Large	809, 810	Tennessee Eastman	Kingsport, TN			2.328
OINC-Large	711	Chevron Chemical Co.	Belle Chasse, LA		x	3.314
OINC-Large	705, 490	Ciba-Geigy Corporation	McIntosh, AL			2.328
OINC-Large	353, 354	Dow Chemical Co.,	Midland, MI			2.328
OINC-Large[b]	334	3M	Cottage Grove, MN		x	1.197
OINC-Large	600	Dow Chemical	Freeport, TX		x	2.489
OINC-Large	B20	GSX Chemical Services	Cleveland, OH			2.083
OINC-Large	806	Amoco Oil, Co.	Whiting, IN			2.328
OINC-Large	483	Hoechst Celanese	Seabrook, TX			2.522
OINC-Large	A50	Quantum Chemical Company	La Porte, TX			3.243
OINC-Large	477, 478, 805	American Cyanamid	Hannibal, MO			2.328

(continued)

Table 4-3. (continued)

Site Type	Site IDs	Company Name	Location	Area	WHB	Adjusted Facility Sampling Weight[a]
OINC-Small	A31	Hercules, Inc	Franklin, VA		x	0.981
OINC-Small	A26	Eastman Chemical Co,	Magness, AR		x	2.026
OINC-Small	B32	Miles Corp.	Baytown, TX			3.965
OINC-Small	A14	Basf Corporation	Geismar, LA			3.049
OINC-Small	A46	OSI Specialties Inc	Sisterville, WV			3.965
OINC-Small	824	Penwalt Corp.	Thorofare, NJ			3.965
OINC-Small	A47	Phillips Research Center	Bartlesville, OK			3.965
OINC-Small	B37	Pine Bluff Arsenal	Pine Bluff, AR	x		7.236
OINC-Small	340	Miles Inc.	New Martinsville, WV		x	1.319
OINC-Small	704	Ashland Chemical Company	Los Angeles, CA		x	1.138
OINC-Small	701	Eli Lilly and Company	Clinton, IN			3.965
OINC-Small	708	Burroughs Welcome	Greenville, NC			3.965
OINC-Small	A55	Schenectady International, Inc.	Rotterdam Jct., NY		x	2.026
OINC-Small	B44	Shell Chemical Co.	Deer Park, TX			2.847
OINC-Small	453	Cargill Chemical Products	Forest Park, GA		x	2.026
OINC-Small	906	Monsanto Agricultural Company	Muscatine, IA			3.965
OINC-Small	904	First Chemical Co.	Pascagoula, MS		x	1.319

(continued)

Table 4-3. (continued)

Site Type	Site IDs	Company Name	Location	Area	WHB	Adjusted Facility Sampling Weight[a]
OINC-Small	468	Lonza Chemical	Conshohocken, PA			3.965
OINC-Small	A45	Occidental Chemical Vcm	Deer Park, TX			2.847
OINC-Small	B23	Huntsman Corp.	Port Neches, TX			3.049
OINC-Small	B18	Georgia Gulf Corp	Plaquemine, LA		x	2.026
OINC-Small	B31	Merck and Co.	West Point, PA			3.049
OINC-Small	342	Upjohn Company	Kalamazoo, MI		x	1.138
OINC-Small	725	Zeneca	Bayonne, NJ			3.965
OINC-Small	493, 494	U.S. Army Tooele Depot North	Tooele, UT	x		7.236

CINC = Commercial incinerator.
CK = Cement kiln.
LWAK = Lightweight aggregate kiln.
OINC = On-site incinerator.
WHB = Waste heat boiler.

[a]These facility weights do not include the sector-level population component.

[b]Facilities modeled for proposed rule.

4.1.6 Analysis Weights

This section discusses how the analysis weights and their components were calculated. The analysis weights were used to make inferences about individual and population risk estimates from the modeled facilities to all HWC facilities. Analysis weights were derived separately for each of the modeled facilities. These weights were then applied to each of the sector-specific risk estimates to create weighted estimates, which could then be used to create cumulative risk distributions for a given combustor category. The overall analysis weight was calculated as the product of two weight components: (1) facility sampling weight, including facility poststratification adjustments, and (2) sector-specific population weight. Each of these weight components is described below.

4.1.6.1 <u>Facility Sampling Weight</u>. The facility sampling weight (WT1) for each sampled facility was the reciprocal of the probability of selection. In most cases, the probability of selection was simply the stratum sample size divided by the stratum frame size. However, the inclusion of a supplemental strata (i.e., the waste heat boilers) complicated the probability structure and resulted in some facilities having multiple chances of selection. Hence, the facility probability of selection was not uniform within a given combustor category and is defined as:

$$\pi_h(i) = \begin{cases} 1 & \text{for certainty facilities, else} \\ \dfrac{n_h}{N_h} & \text{for facilities with one selection chance, else} \\ P_1 + (1-P_1)\,P_2 & \text{for facilities with two selection chances, else} \\ P_1 + (1-P_1)\,P_2 + (1-P_1)\,(1-P_2)\,P_3 & \text{for facilities with three selection chances,} \end{cases} \qquad (4\text{-}2)$$

where

h	=	sampling stratum
P_1	=	probability selected in first possible stratum
P_2	=	probability selected in second possible stratum
P_3	=	probability selected in third possible stratum.

Therefore, the facility sampling weight was assigned as follows:

$$WT1 = 1 / \ \pi_h(i) \ . \qquad (4\text{-}3)$$

Table 4-4 lists the possible selection strata for the facilities with multiple chances of selection and indicates how classification changes affected the possible selection strata.

Table 4-4. Facilities with Multiple Chances of Selection

Facility IDs	Actual Selection Stratum	Possible Selection Strata	Classification Change
209, 324, 359, 601	4	4, 7, 8	None
a31	5	5, 7, 8	OINC-L ⇒ OINC-S
342, 704	6	6, 7, 8	none
b20	6	6, 7, 9	OINC-S, WHB ⇒ OINC-L, non-WHB
b31	6	6, 7	WHB ⇒ non-WHB
a14, b23	7	6, 7	WHB ⇒ non-WHB
600	8	5, 8	None
340, 904	8	6, 8	None
463, stg	9	5, 9	None
a32	9	6, 9	OINC-S ⇒ OINC-L

OINC-L = On-site incinerators - large.
OINC-S = On-site incinerators - small.
WHB = Waste heat boilers.

Facility Poststratification Adjustment. The cumulative design modifications (described in Section 4.1.5) have the effect of reducing the efficiency of the sample. To improve the quality of the sample estimates, the facility sampling weights (WT1) were adjusted to force sample estimates of the total number of facilities in the categories listed in Table 4-5 to equal the known totals for these categories. The categories were established by cross-classifying combustor type with waste heat boiler status and combustor type again with area source status.

The individual facility adjustment factors are the quantities λ_i in the equation

$$\sum_{i \in S} w_i \, \lambda_i \, \underset{\sim}{x}_i' = \underset{\sim}{T}', \qquad (4\text{-}4)$$

where the range of summation is taken over all facilities in the sample and

w_i = facility sampling weight (i.e., WT1 defined above)

$\underset{\sim}{x}_i'$ = transpose of a vector of indicator (0,1) variables identifying the categories of facilities listed in Table 4-5

$\underset{\sim}{T}'$ = transpose of the vector of known category totals.

Table 4-5. Average Weight Adjustment Factors from Exponential Model for Poststratifying to Facility Population Totals

Exponential Model Variable	Population Control Total	Average Facility Sampling Weight Adjustment Factor
Combustor Type / Waste Heat Boiler Status		
Cement Kiln	18	1.00
Lightweight Aggregate Kiln	5	1.00
Commercial Incinerator, Waste Heat Boiler	8	1.55
Commercial Incinerator, not Waste Heat Boiler	12	1.01
Large On-site Incinerator, Waste Heat Boiler	7	1.20
Large On-site Incinerator, not Waste Heat Boiler	36	0.84
Small On-site Incinerator, Waste Heat Boiler	14	0.53
Small On-site Incinerator, not Waste Heat Boiler	65	1.13
Combustor Type / Area Source Status		
Cement Kiln, Area Source	2	0.87
Cement Kiln, not Area Source	16	1.02
Lightweight Aggregate Kiln	5	1.00
Incinerator, Area Source	28	1.59
Incinerator, Not Area Source	114	0.85

The adjustment factors were computed as the solutions to the exponential regression relation

$$\lambda_i = \exp\left(\alpha + \underset{\sim}{x}_i \underset{\sim}{\beta}\right), \tag{4-5}$$

where

α = value of the relation at $\underset{\sim}{x}_i = \underset{\sim}{0}$ (i.e., the intercept)

$\underset{\sim}{\beta}$ = vector of regression coefficients relating the weighted sample observations to the facility categories.

The α- and β-values were determined numerically to satisfy Equation 4-5. The solutions were constrained so that $0.5 \leq \lambda_i \leq 2.0$. The imposition of these constraints ensured that sampling variances were not excessively inflated because of unequal weighting effects associated with making the (poststratification) adjustments.

The adjusted facility sampling weights are the product of the initial facility sampling weights (WT1) and the adjustment factors (λ_i). The adjusted facility sampling weights are presented in Table 4-3. The average weight adjustment factors and the known population counts (control totals) are shown in Table 4-5 for each of the defined combustor categories.

4.1.6.2 Sector-Population Weight (WT2). Since all 16 sectors for every sampled facility were selected (i.e., included in the risk characterization), the sector sampling weight is 1.0. However, because the analysis is at the sector level and estimates on the human population are being made, the sector weight needs to be multiplied by the human population in each sector. Consequently, the sector population weight is

$$WT2 = 1 \bullet pop_{ij} \qquad (4\text{-}6)$$

where

i = facility
j = sector.

For recreational fishers, subsistence farmers, and subsistence fishers, the human population was set to 1 because information was not obtained to approximate sector-level populations for those groups (i.e., these receptor populations were not weighted).

4.1.7 Variance Estimation (Confidence Intervals)

Most statistical software packages assume simple random sampling from an infinite population and are not appropriate for variance estimation of sample survey estimates. That is, they do not compensate for survey design features such as stratification, clustering, and sampling from a finite population. Hence, they would produce biased variance estimates for sample survey data. To account for these survey design features, all of which are components of the HWC risk analysis, the majority of risk estimates (and associated confidence intervals) for the HWC risk analysis were computed using RTI's statistical software package, SUDAAN®. SUDAAN is a multiprocedure package that takes into account survey design features (i.e., sample design parameters can be specified and correct standard errors can be computed).

In addition, for probability-based sample surveys, most estimates are nonlinear statistics. Hence, the variances of the estimates cannot be expressed in closed form. For example, a mean or proportion, which is expressed as $\Sigma wy / \Sigma w$, is nonlinear because the denominator is a survey estimate of the (unknown) population total. SUDAAN offers both the Taylor series linearization and replication methods (BRR and Jackknife) for robust variance estimation of nonlinear statistics. For this analysis, the Taylor series linearization method was used. This method computes the Taylor series approximation of the nonlinear statistic and then substitutes the linear representation into the appropriate sample design variance formula.

There were four basic types of estimates computed by SUDAAN for the HWC risk analysis:

- Cumulative distributions for risks (or hazard quotients) not weighted by population
- Population-weighted individual risk (or hazard quotient) percentiles
- Population estimates of cancer incidence (both local and national)
- Proportion of population with risk (or hazard quotient) greater than the health benchmark level.

The uncertainty of all the estimates was measured by 90 percent confidence intervals. The confidence intervals for the percentiles were computed internally by SUDAAN. To obtain confidence intervals of a given percentile, SUDAAN first computes the confidence intervals for the cumulative distribution based on the sampling error of the cumulative distribution. Then, the confidence bounds for a given percentile are determined from the confidence bound formulas of the cumulative distribution. This method was used to compute confidence intervals for cumulative distributions for risks (or hazard quotients) not weighted by population and population-weighted individual risk (or hazard quotient) percentiles.

The 90 percent confidence intervals for the population estimates of cancer incidence (both local and national) were computed from a log transformation. Because the cancer incidence estimates are small and the sample sizes are small for some domains, the underlying distribution was assumed to be asymmetric and the log transformation was used to compute asymmetric confidence intervals. These asymmetric intervals are more balanced with respect to the probability that the interval covers the true population value than do standard symmetric confidence intervals. For this analysis, only 90 percent confidence intervals were calculated. To illustrate the method, let

T = estimated population total ($\Sigma w_i\, x_i$)
L = natural log of T
SE(L) = standard error of L.

The 90 percent confidence intervals for L were then calculated as

$$A = L - t_{.05}SE\{L\}$$
$$B = L + t_{.05}SE\{L\}. \qquad (4\text{-}7)$$

The Student's t-distribution with 70 degrees of freedom was used instead of assuming a normal distribution. However, with 70 degrees of freedom, the normal and Student's distributions are essentially equal. The degrees of freedom are equal to the number of selected facilities (76) minus the number of analysis strata (6), which are the first six strata listed in Table 4-2.

By taking the exponential values of A and B, the confidence intervals for the population total, T, are

$$T_{lower} = \exp(A)$$
$$T_{upper} = \exp(B) \ . \tag{4-8}$$

For the proportion of population with risk (or hazard quotient) greater than the health benchmark level, the logit transformation, ln[p/(1-p)], was used to compute the confidence intervals. The confidence intervals using the logit transformation were computed in a manner similar to that used for the log transformation for the population totals. The logit transformation prevents estimates of prevalence rates from being either less than zero or greater than unity. The transformation itself is given by

$$X_d = \mathrm{Logit}\left\{\hat{P}_d\right\} = \ln\left\{\frac{\hat{P}_d}{1 - \hat{P}_d}\right\} \tag{4-9}$$

where $\hat{P}_d$ is the estimated prevalence rate for the *d*-th reporting domain (e.g., type of chemical by receptor population). The interval estimate can be written as

$$\mathrm{Prob}\left\{P_{d,\ell} \le P_d \le P_{d,u}\right\} = 1 - \alpha \, . \tag{4-10}$$

On the transformed scale, the interval estimate becomes

$$\hat{X}_d \pm t_{\alpha/2}\,\mathrm{SE}\left\{\hat{X}_d\right\} = \hat{X}_d \pm t_{\alpha/2}\left(\frac{\sqrt{\mathrm{Var}\left\{\hat{P}_d\right\}}}{\hat{P}_d\left(1 - \hat{P}_d\right)}\right), \tag{4-11}$$

and the inverse transformations

$$P_{d,\ell} = \frac{1}{1 + \exp\left\{\hat{X}_d - t_\alpha\,\mathrm{SE}\left\{\hat{X}_d\right\}\right\}},$$

$$P_{d,u} = \frac{1}{1 + \exp\left\{\hat{X}_d + t_\alpha\,\mathrm{SE}\left\{\hat{X}_d\right\}\right\}} \tag{4-12}$$

provide the upper and lower bounds of the intervals on the arithmetic scale. The intervals in this case are not necessarily symmetric.

Confidence intervals generated for the HWC risk analysis are symmetric around each of the quantiles being considered and are calculated as follows. Denote the distribution function of interest by

$$F(x) = \frac{1}{N}\sum_{g=1}^{N} I(y_g \le x) \tag{4-13}$$

where the subscript $g = 1, 2, \ldots, N$ identifies units in the population, in this case sectors, and

y_g = value returned by the risk model for the g-th sector,

$$I(y_g \le x) = 1, \text{ if } y_g \le x,$$
$$= 0, \text{ otherwise.}$$

The quantiles of the distribution are defined by the values k such that

$$F(x_k) \le Q_p \le F(x_{k+1}). \tag{4-14}$$

SUDAAN estimates the distribution function by

$$\hat{F}(x) = \frac{\sum_{g \in S} w_g I(y_g \le x)}{\sum_{g \in S} w_g}, \tag{4-15}$$

where w_g are the sampling weights, and finds the values $k = 1, 2, \ldots, p$ such that $\hat{F}(x_k) \le Q_p \le \hat{F}(x_{k+1})$. The confidence interval is computed using the standard error

$$SE\{Q_p\} = \frac{\hat{U}_p - \hat{L}_p}{2t_{\alpha/2}} \tag{4-16}$$

where $\hat{U}_p$ and $\hat{L}_p$ are the limits implied by $\hat{F}(x_k) \pm t_{\alpha/2} SE\{\hat{F}(x_k)\}$, and $t_{\alpha/2}$ = value of Student's t-distribution at the significance level $\alpha/2$.

Hence, the variance of interest, that is the quantity $(SE\{Q_p\})^2$, involves the point on the estimated distribution function $\hat{F}(x_k)$ and the standard error associated with the estimate at that point. The intervals are seen to be symmetric about the quantile. For some of the risk (and HQ) percentiles of cumulative distributions, confidence intervals could not be generated because of an insufficient sample size or insufficient spread of modeled risk values.

4.2 Facility Operating Characteristics and Emissions Estimates

This section describes the facility-specific engineering and annual emissions data used to conduct air modeling for purposes of generating sector-level air concentration and deposition estimates for modeled HWC facilities[5]. Assumptions concerning operational, facility-specific engineering and annual emissions estimates are presented.

Section 4.2.1 describes the database used to characterize modeled HWC facilities and Section 4.2.2 describes operating scenarios, engineering data, and annual emission estimates.

4.2.1 Facility Database

In conducting the HWC risk analysis, information on the universe of facilities as well as HWC facility-specific engineering data were required. A brief overview of the data sources and methodologies used to develop these data is presented here. A more comprehensive discussion is provided in U.S. EPA (1999b).

The database used in this analysis contains the following facility-specific data:

- Facility equipment and operational data (e.g., engineering data including stack heights, combustors, air pollutant control device [APCD], temperatures, exit velocities)

- Emission rates for constituents discharged to the atmosphere (e.g., metals, chlorine, PM, PCDD/PCDF, CO, and HCl) from the facility's main stack.

The HWC facilities included in the database are all facilities known to be operational in 1997.

These data were revised since proposal in an effort to incorporate additional facility-specific information as it became available and to address data issues raised in public comments. Specifically, the database was augmented with facility-specific information obtained during an initial comment period (at proposal), a subsequent Notice of Data Availability (NODA) comment period, and further data-gathering efforts involving visits to Regional EPA and state environmental offices, which were conducted in the fall of 1997.

EPA published a notice in the *Federal Register* covering the database that was used to set the floor levels via a NODA on January 7, 1997. The database contained all the information available from trial burn and certificate of compliance reports that was used in the analysis, including emissions data and engineering information on APCD and operating parameters as well as stack information. This information was used to characterize stack emissions where measurements were available and for imputing exhaust gas concentrations where they were not.

[5] The term "engineering data," as used here, refers to data used to characterize the physical release of emissions (e.g., stack height, stack diameter, and exit velocity), including all parameters necessary for conducting air modeling.

Those facilities that were in the universe of facilities covered by the rule but for which EPA did not have test reports were not in the database.

This additional information, obtained chiefly from newly identified trial burn and compliance test reports, resulted in adjustments to facility-specific engineering parameters and emissions estimates. In some cases, the new information resulted in a facility being removed from the database (e.g., closure) or changing its classification from one source category to another (e.g., OINC-L with WHBs to OINC-L without WHBs).

4.2.2 Facility-Specific Engineering and Emissions Data

The HWC facilities modeled for this risk analysis may contain one or more combustion units with associated stacks. Although risks were assessed at the facility level, air modeling was conducted separately for each stack. Emissions that result from materials handling, fugitive releases, emergency safety valve releases, disruptions in the normal combustion operation, startup, and shutdown (none of which are subject to the MACT standards) were not modeled. Therefore, emission rates may not be representative of all operating scenarios experienced at an actual facility.

These combustion units were assumed to be operating continuously for 24 h/d, 365 d/yr. Annual stack emissions were calculated assuming continuous operation for an entire year. Therefore, annual emissions rates may not be representative of actual facility operation with regard to temporal fluctuation in emissions and actual times of emission release. Short-term emissions derived from annual emission rates and used in air dispersion and deposition modeling were based on 8,760 hours of operation per year (see Section 5.1 for more information on air modeling inputs).

Emissions scenarios were developed for a base case and for three regulatory alternatives:

- **Baseline** —The baseline scenario assumes emissions rates associated with normal operation as they currently exist without application of additional air pollution controls.

- **MACT-Standard**—The MACT standard scenario assumes emissions rates based on a set of air pollution controls that would be required to satisfy the final rule. These controls are a mixture of MACT floor and MACT beyond-the-floor requirements.

- **MACT-Floor**—The MACT floor scenario assumes emission rates based on a set of air pollution controls required to satisfy the minimum control requirements for HWC facilities under Section 112(d)(3) of the CAA.

- **MACT-Beyond-the-Floor—**The MACT beyond-the-floor scenario assumes emission rates based on a set of controls necessary to achieve a greater degree of emissions reduction than is required for the MACT floor scenario. These more effective controls are applied to dioxins, mercury, lead, hydrogen chloride, and chlorine gases for certain combustor categories.

No increase in emissions was assumed for those facilities that were operating below the design level needed to satisfy the MACT standard. That is, emission rates used in this risk assessment do not reflect that a facility would make changes to their operations and increase to an emissions level higher than they were emitting before the standards. For a more detailed discussion of the regulatory scenarios evaluated for the HWC MACT rulemaking, see U.S. EPA (1999b).

Engineering data were required to estimate emissions and as input to air modeling. The following categories of facility-specific engineering data were used for air modeling: stack location (latitude and longitude), stack height (m), stack inside diameter (m), exit velocity (m/s), stack gas temperature (K), and building height and width (m). Facility-specific engineering data used for air dispersion modeling are presented in Appendix B.

The final list of constituents selected for evaluation in the HWC risk analysis consisted of

- 17 dioxin/furan congeners
- 14 metals
- Chlorine and hydrogen chloride
- $PM_{2.5}$ and PM_{10}.

Emissions estimates were made for all chemical constituents covered by the rule for which sufficient data were available. These included chlorine-substituted dibenzo(*p*) dioxins and dibenzofurans, elemental mercury (Hg^0), divalent mercury (Hg^{+2}), lead, cadmium, arsenic, beryllium, trivalent chromium (Cr^{+3}), hexavalent chromium (Cr^{+6}), chlorine, and hydrogen chloride. In addition, emissions estimates were made for particulate matter (PM_{10} and $PM_{2.5}$) and nine other metals, three of which (cobalt, copper, and manganese) were not assessed at proposal but were included in the risk assessment for the final rule. Chemical-specific emissions estimates were not made for organic constituents other than dioxins and furans (e.g., various products of incomplete combustion) due to insufficient emissions measurement data. Risks from all constituents for which chemical-specific emissions estimates could be made as well as from PM were evaluated in this risk assessment.

The original facility-specific emissions concentration and flow rate data were obtained primarily from trial burn and certificate of compliance test reports. When more than one source of emissions data was available, data were obtained from the report based on the most recent sampling. An imputation scheme was used to fill in missing emissions data for HWC facilities. In conducting imputation, efforts were made to match the missing data to the group of facilities from which values were being imputed based on similarities in equipment and operations (i.e., data would be imputed for a given facility from a set of facilities with characteristics similar to those of that facility). Facilities were matched for purposes of imputation to improve the representativeness of the imputed data. An in-depth discussion of the imputation procedure as well as the overall approach used in developing the database is provided in U.S. EPA (1999b).

4.3 Site Characterization for Modeled Facilities

This section describes the methodologies and data sources used in site characterization specifically with regard to

- # Selecting and characterizing waterbody/watersheds for inclusion in risk modeling

- # Establishing site-specific human and livestock populations.

Site characteristics associated with air dispersion modeling (e.g., terrain, meteorology) are discussed in Section 5.1.

4.3.1 Study Area

The HWC risk analysis conducted for the final rule generates spatially refined human health and ecological risk results based on a 16-sector study area template. To achieve the desired degree of spatial resolution for this risk assessment, a 20-km radius polar grid was used (see Figure 4-2). This polar grid, which is centered on the geographic coordinate for the HWC facility, was divided into 16 sectors and numbered as indicated in Figure 4-2. An individual polar grid, together with its HWC facility, is termed a "study area." The term "sector" refers to the 16-sector grid that defines the study area. The sector polygons were created by dividing four concentric circles around the site location (2, 5, 10, and 20 km radius) by the north-south and east-west axes.

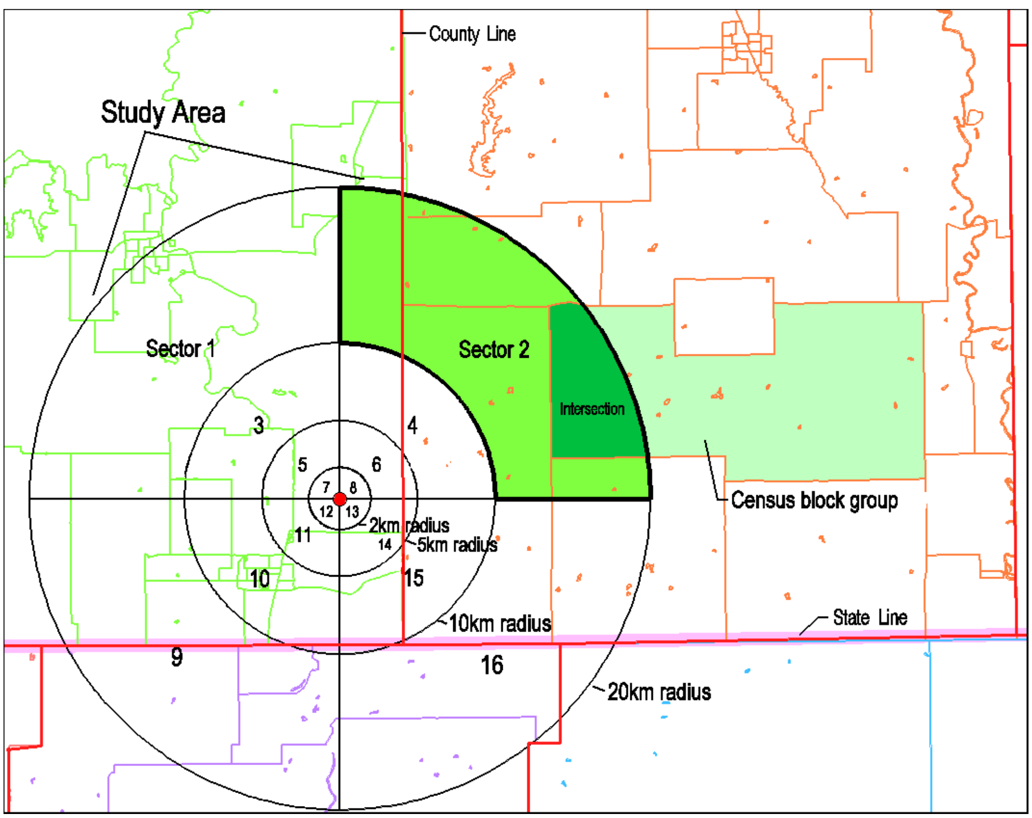

Figure 4-2. Example of study area, sectors, and area-weighted population apportionment.

Population counts for each receptor population at the sector level were combined with individual sector-level risk estimates to determine human health risks within study areas. Similarly, livestock population counts at the sector level were combined with sector-level dioxin concentration estimates to estimate national population risks resulting from exposure to dioxin contained in livestock raised within study areas.

Because of the volume of data required for the analysis, semiautomated techniques were used to access existing nationwide databases to provide the site-specific data used in site characterization. A geographic information system provided the platform for projecting the impact of HWC emissions on individual study areas and watersheds/waterbodies and for characterizing land use for estimation of human and ecological risk (e.g., location, shape, and size of watersheds and waterbodies and densities of human and livestock populations). The human health component of the HWC risk assessment includes risk estimates for receptor populations located within the modeled study areas (e.g., beef cattle farmers and recreational fishers). The HWC risk assessment also includes risk estimates for those human populations located outside of modeled study areas that may be impacted by ingestion of dioxin contained in food commodities produced within these study areas (e.g., individuals eating beef from beef cattle that were raised within a given study area and thereby exposed to dioxin from the associated HWC facility).

A GIS was selected as the platform for conducting the site characterization component of the HWC risk analysis because it can be easily automated and can perform spatial overlay of georegistered data. Most of the GIS processing was conducted using ARC/INFO for UNIX workstations; some took place in the PC environment with ARC/VIEW. The term “program” is used throughout this section to refer to Arc Macro Language (AML) scripts, a batch-process scripting language used with the ARC/INFO GIS software. The term “coverage” refers to a GIS map layer (e.g., geographically referenced digital points, lines, or polygons with attached data).

The GIS modeling results provided three sets of data inputs for the risk analysis:

- Waterbody characteristics and average air concentration and deposition values by watershed and waterbody within the study area

- Average air concentration and deposition values by sector within the study area

- Spatially averaged human and livestock populations by sector.

The remainder of this section discusses the various methodologies used to derive these data inputs for the risk analysis.

4.3.2 Waterbody/Watershed Selection, Delineation, and Characterization

With the exception of one site, from one to four waterbodies were selected for inclusion in the HWC risk analysis from each study area. For one site and region, there were no waterbodies modeled. Selected waterbodies were delineated and characterized. These waterbodies, termed “modeled waterbodies,” were used to provide site-specific data used in the risk analysis. In characterizing modeled waterbodies, the following attributes were compiled for each waterbody and associated watershed:

- Watershed area
- Length of stream reach
- Waterbody area
- Universal soil loss equation (USLE) parameter
- Flow velocity and discharge
- Stream width/depth
- Total suspended solids concentrations.

Each of these attributes was defined for that portion of the watershed/waterbody located within the 20-km radius study area under consideration.

A combination of desktop evaluations using available maps/databases and GIS techniques was used to obtain site-specific values for each of these attributes. This section describes the approach used to select and delineate modeled waterbodies. In addition, the data sources and methodologies used in site-specific characterization of those modeled waterbodies are described.

4.3.2.1 Compilation of Study Area Data. Existing data layers were compiled to create a single comprehensive map for each study area. These maps, which were generated with GIS tools, are called "compilation maps." They were used to select waterbodies for inclusion in the study and delineate their associated watersheds. These 17 x 17 inch color maps were generated using an automated batch script that started with the point coverage of the site's location and then added the following map layers:

- **Sector boundaries:** Generated previously with an automated batch script

- **RF3 data:** EPA stream reach files (U.S. EPA, 1994) generated from 1:100,000 scale U.S. Geological Survey (USGS) digital line graphs (DLGs)

- **Drinking Water Supply Sites:** Supplied from the BASINS CD-ROM database (Laveck and Coombs, 1996)

- **Stream Gaging Stations:** Obtained from the BASINS CD-ROM and WATSTORE databases (USGS, 1994)

- **Pseudo drainage basin lines:** Generated from 1:250,000 Digital Elevation Model (DEM) coverages obtained from USGS. (A DEM consists of an array of elevation values for ground positions that are usually at regularly spaced intervals.)

4.3.2.2 Waterbody Selection. The following criteria were used to select modeled waterbodies/watersheds for inclusion in the HWC risk analysis:

- # **Probable impact from facility emissions:** Those waterbodies located in the direction of prevalent winds and relatively close to the HWC facility were favored in selecting waterbodies for modeling to ensure that risks generated for study areas included more heavily impacted waterbodies.

- # **Probable recreational use (including fishing):** Although it is difficult to determine patterns of recreational use at waterbodies from the maps used in selecting modeled waterbodies, characteristics suggestive of recreational use (e.g., larger waterbody size, location in favorable land-use areas, and good public access as determined from road and parking lot patterns) were considered in selecting waterbodies for inclusion in the HWC risk analysis.

- # **Drinking water source:** Priority was given to waterbodies identified as drinking water sources. If several drinking water sources were identified for a given study area, priority was given to the one likely to be impacted to a greater extent by HWC emissions due to its location.

In general, the waterbodies selected for modeling favor those located in areas more heavily impacted by HWC emissions and do not represent a random sample that can be considered representative of all waterbodies located across the study areas. There is, however, an important caveat to this general statement. In selecting waterbodies for a given study area, often a different waterbody was selected to match each of the three criteria listed above (e.g., waterbody A may be selected from a more impacted location within the study area, waterbody B may be selected because it looked like a probable recreational location, and waterbody C may be selected because it was the drinking water source closest to the facility). Because all three criteria were considered in selecting waterbodies for inclusion in the HWC risk analysis, the waterbodies that were selected do not always represent those waterbodies most impacted by HWC emissions. In certain instances, the goal of including a drinking water source or a waterbody that appeared to be a likely location of recreational activity resulted in the exclusion of a more heavily impacted waterbody.

4.3.2.3 <u>Watershed Delineation</u>. The compilation maps and USGS 7.5 minute 1:24,000 scale quadrangle maps were used to delineate the watersheds for the selected waterbodies. The following delineation protocol was applied to each selected waterbody:

- # Watershed boundaries were delineated by starting at the farthest downstream point of the selected stream (or outlet of the selected lake) that was still within the 20-km radius. A line was then drawn perpendicular to the topographic contour lines upgradient from that point. This line was extended until it reached the point at which the elevation ceased to increase or until it intersected with the boundary of the study area. Then, starting again from the farthest downstream point, a line was drawn (in the opposite direction) perpendicular to the contour lines until the elevation ceased to increase. The endpoints of these two lines were connected through the peaks and ridges on the map or along the boundary of the study area, whichever covered less area.

- # Only the watershed area that drains into the selected waterbody **before** the waterbody flows out of the 20-km study area was included. If a tributary to the selected waterbody merged with the waterbody downstream of the study area, the area that drained to this tributary was not included in the watershed area.

- # No watershed area that lies outside of the radius was included. Because only the 20-km study area was of interest, the radius acts as a cutoff distance for watershed delineation. Cutting off the watershed at 20 km is consistent with the definition of the study area used for this risk assessment. The result is that contaminant loading to waterbodies is based only on that portion of total emissions that deposit on watershed areas located in the study area. As noted below, however, waterbody flow is based on total flow (including tributaries located outside the study area), which best characterizes the waterbody's properties. The resulting uncertainty generally underestimates that portion of waterbody constituent concentrations that results from watershed loadings.

- # Only the main stem for the selected streams (i.e., no tributaries) and waterbody surface areas for lakes and reservoirs were included.

- # Arcs (lines) and polygons were digitized manually with a standard digitizing tablet and ARC/INFO workstation. Lakes and watershed polygons were labeled with name and site identification. Stream/river lines were labeled with name, width, and site identification. Stream coverages were processed in a program that changed the line coverage into a polygon coverage based on each stream's width.

Watershed delineation and digitization allowed for collection of the necessary model input parameters. Watershed area, stream length, and waterbody area values were extracted from the data tables associated with the digitized coverages.

Quality control measures were taken on each major step of the delineation process. A quality control check was completed after manual delineation to ensure correct watershed delineation and on the completed GIS coverages to ensure correct translation of watershed area and other parameters.

4.3.2.4 <u>Watershed Universal Soil Loss Equation Parameters</u>. The Indirect Exposure Emissions Model used for this risk analysis uses the USLE to estimate soil erosion losses (X_e) from watersheds that drain into modeled waterbodies surrounding each hazardous waste combustor site (Section 5.3.2). USLE is an empirically derived equation originally developed by the Soil Conservation Service (SCS) of the U.S. Department of Agriculture (USDA) to estimate soil erosion losses from agricultural fields during soil conservation planning. In the IEM methodology, USLE is applied in the context of the Gross Erosion Sediment-Delivery Ratio method outlined in USDA (1978) and described in greater detail in the SCS *National Engineering Handbook* (USDA, 1971). Gross erosion is defined as the summation of erosion from all sources within a watershed, as estimated for sheet and rill erosion by USLE. The sediment delivery ratio adjusts gross erosion rates to account for eroded soil that does not reach the waterbody in question. USLE requires inputs to estimate soil erosion losses, including rainfall and runoff factor (R), soil erodibility factor (K), topographic factor (LS), cover and management factor (C), and supporting practice factor (P). In this context, USDA (1978) suggests the use of watershed-averaged values for K, LS, C, and P to simplify computational and data collection requirements. With the exception of K, this approach was adopted for developing site-specific USLE gross erosion loss estimates for the combustor sites. Each of these inputs is discussed below.

Rainfall and Runoff Factor. This factor quantifies the rainfall impact effect and provides relative information on the amount and rate of runoff associated with rain. The rainfall and runoff factor is the number of rainfall erosion index units, plus a factor for snowmelt or applied water where such runoff is significant. The rainfall erosion index for a given rainfall event is equal to the total storm energy times the maximum 30-min intensity. Local values of the erosion index were taken directly from the isoerodant maps provided in USDA (1978) by locating each site based on its location in the United States and selecting the closest value, interpolated as necessary. The rainfall erosion index, however, does not account for runoff associated with surface thaws and snowmelt. Soil erosion by thaw runoff is most pronounced in the northwest United States, but it may be significant in other northern states. In this analysis, surface thaw and snowmelt were not considered, which for some facilities could underpredict the amount of runoff, but the overall influence on not including this information was not expected to be substantial because only a subset of facilities would be affected (none are located in the Pacific Northwest for instance).

Soil Erodibility Factor. This factor accounts for variability in different soils' tendencies to erode. A national value for K for silt loam, which is a predominant soil type both nationally and for the combustor sites, was obtained for consistency with the national parameterization of other soil properties required for the model. STATSGO national soils data, compiled by the STATSGO map unit in the USSOILS database, were used to estimate various central tendency statistics for the more than 1,400 STATSGO map units across the country with silt loam soils and nonzero K values. Results are shown in Table 4-6. All central tendency statistics (mean, median, mode, area-weighted mean) were 0.34, which was the K value used in the analysis.

Table 4-6. National Central-Tendency Statistics: USLE Erodibility Factor (K) for Silt Loam Soils

Statistic	USLE K
Median	0.3400
Area-weighted average	0.3436
Mode	0.3400
Mean	0.3420

Data source: STATSGO/USSOILS national soils database.

Use of a single national value of K does not create any additional uncertainties beyond those associated with assuming a national soil type for all of the other soil properties (e.g., bulk density, porosity). Because soil erodibility is correlated with soil type, there would be a disconnect in using a site-based K and national parameterization for all the other soil properties (i.e., assuming different soils in different model components). The length-slope, cover and management, supporting practice, and rainfall and runoff factors are not strongly correlated with soil type, enabling use of site-specific values for those parameters.

Topographic Factor. The topographic factor quantifies the effects of slope (S) and slope length (L) on soil erosion loss. Both average slope (S) and average slope length (L) are required to determine an average watershed length-slope factor (LS). The STATSGO and USSOILS databases, which are maintained by the USGS, provide spatial information on soil series by map units, spatially contiguous areas with similar soil properties. For each watershed, S was queried from the USSOILS version of the STATSGO database. However, the STATSGO/USSOILS database does not contain data characterizing the slope length component (L) of the LS parameter. Although options were identified for estimating site-specific length values based on watershed area and total stream length, adequate site-specific data on stream length could not be obtained (existing sources such as Reach File Version 3 [RF3] do not contain true first-order streams and, therefore, underestimate total stream length).

Because no consistent national data sources were available for this parameter (other than direct field measurements), national default L values, dependent on slope length, were obtained from personnel at the USDA Natural Resources Conservation Service experienced in erosion prediction, agronomy, and soil services. These L and corresponding S values were estimated for national use in pesticide and construction erosion studies (Weesies, 1998) as shown in Table 4-7.

The site-specific average watershed S values were used to determine the corresponding L values, per Table 4-7. The L values thus determined were used to calculate an average LS for each watershed using the following equation (USDA, 1978):

$$LS = \left(\frac{L}{72.6}\right)^m \bullet (65.41\sin^2\theta + 4.56\sin\theta + 0.065) \tag{4-17}$$

where

L = slope length (feet)

m = exponent dependent on slope (0.5 if S is 5% or more, 0.4 at 3.5 to 4.5%, 0.3 at 1 to 3%, 0.2 at less than 1%).

θ = angle of slope

This application of a site-specific value, a lookup table, and the equation to derive LS results in the use of an LS value that reflects site characteristics yet is not purely site-specific.

Table 4-7. Default Slope Length (L) used in Combustor Risk Analysis

Slope (%)	Default length (ft)	Slope (%)	Default length (ft)
< 0.05	100	9	125
1	200	10	120
2	300	11	110
3	200	12	100
4	180	13	90
5	160	14	80
6	150	15	70
7	140	17	60
8	130	>17	70

Source: Weesies (1998).

Cover and Management and Supporting Practice Factors. The cover and management factor measures the effect of land cover (e.g., crops, forests) on soil erosion losses; the supporting practice factor accounts for erosion control measures that may be applied (such as contour plowing for cropland). The cover management and supporting practice factors were derived from site-specific land use data obtained from GIRAS (Geographic Information Retrieval and Analysis System) data, which are coded using the Anderson land use II classification scheme
(Anderson et al., 1976)[6]. Using the crosswalk shown in Table 4-8, a database was created to convert Anderson codes to broader land use categories for which typical C and P values are available (Wanielista and Yousef, 1993). These values were then spatially averaged in the database to give an average C and P for each watershed.

4.3.2.5 <u>Waterbody Characterization</u>. Parameters that were required for the model but were not supplied by watershed delineation were stream/river velocity, discharge, width, and depth. BASINS Reach Files Version 1 (RF1) was queried by region for the selected waterbodies. Most of the selected waterbodies were listed, with all of the necessary parameter values. Many of the waterbodies had multiple data sets associated with different locations along the waterbody. If there was more than one valid datapoint, the most inclusive data were selected; thus, dilution effects from all tributaries were included.

[6] It should be noted that HWC air modeling runs were also completed with roughness height obtained from GIRAS coverage interpretation (see Section 5.1).

Table 4-8. Cover Factor (C) and Supporting Practice Factor (P) by Land Use Code

Wanielista and Yousef (1993)			
Land Use Type	**C**	**P**	**Anderson Land Use Codes (Anderson et al., 1976)**
Forestland	0.005	1	4 Forest (41-43)
Cropland	0.08	0.5	21 Cropland & Pasture; 25 Other agricultural land
Pastureland	0.01	1	3 Rangeland (31-33); 81-82, 83-84 Tundra
Urban	0.01	1	1 Urban or built-up land (11-17)
Water	0	1	5 Water (51-54); 6 Wetland (61, 62)
No erosion	0	1	74 Bare rock; 91 Snowfields; 92 Glaciers
No cover	1	1	23 Confined feeding operations; 7 Barren land (71, 73, 76); 83 Bare ground

If the waterbody did not appear in the RF1 tables, the BASINS Stream Gaging File was queried. This file provides values only for waterbodies' discharges. The discharge value was then entered into three equations derived from Keup (1985): velocity (ft/s) = $1.0662x^{0.127}$, width (ft) = $5.1867x^{0.4559}$, and depth (ft) = $0.1808x^{0.4171}$ where x (ft^3/s) is the discharge value.

If a discharge value for a particular waterbody was not available in BASINS Stream Gaging File, the USGS WATSTORE database was queried. If a value was available, it was used to estimate the velocity, width, and depth of the waterbody using the three equations from Keup (1985).

If no discharge data were available, the parameter estimations were derived from stream order using RF3 maps or preferably 1:24,000 USGS quad maps. Strahler's stream order classification system was used to order the selected streams. The stream order (1-10) was used in Keup's table (Keup, 1985) to estimate values for discharge, velocity, width, and depth.

Water column concentrations are intrinsically dependent upon the concentration of total suspended solids (TSS). TSS concentration was included in the surface water model as a parameter because there was insufficient information about other parameters (e.g., benthic burial rate for sediments) for both lakes and flowing waterbodies to allow the model to calculate TSS concentrations. Modeling TSS without sufficient information would produce an unacceptable level of uncertainty in the TSS values to which the modeled constituent concentration values are sensitive. Therefore, it was decided to set the parameter value for TSS and treat the benthic burial rate as a variable (Section 5.3.3.1).

Values for TSS were developed from STORET TSS data using a regional approach (http://www.epa.gov/OWOW/STORET/). A regional approach was selected because site-specific TSS estimates could not be calculated reliably with the existing data due to limitations in the STORET data. TSS data collection involved assigning the combustor sites to the regions and establishing typical (or central tendency) TSS values for each region.

The USGS Hydrologic Regions (Seaber et al., 1987) were selected as appropriate regions for the analysis because they are large enough to have an adequate number of STORET TSS values, adequately diverse to show some variability in TSS, and generally accepted as a way of organizing hydrologic and water quality data. Of the 18 continental U.S. Hydrologic Regions, 12 (2-8, 10-12, 16, 18) have modeled HWC facilities.

Ambient water quality monitoring data for TSS were extracted from EPA's mainframe computer located in Research Triangle Park, NC, by hydrologic region for all years of record (1960 to 1997). Data were extracted separately for flowing waterbodies (streams and rivers) and still waterbodies (lakes and reservoirs) because flowing water was anticipated to have significantly different TSS levels.

The STORET data (U.S. EPA, n.d.) (in flat file format) were read into SAS™ for statistical analysis. The SAS PROCUNIVARIATE procedure was used to calculate median TSS values for each region of interest over the entire period of record. The median was selected as the best central tendency statistic because (1) it does not require distributional assumptions and (2) it is relatively stable and not as sensitive as the mean to extreme values that may result from natural variability or errors in the STORET data.

Similarly, a large number of TSS values was needed so that extreme values or data of questionable representativeness would not bias the calculated median. For rivers and streams, thousands and often tens of thousands of TSS values were available in each hydrologic region and robust regional medians could be calculated. For lakes and reservoirs, fewer TSS measurements were available and it was necessary to combine like regions where possible to compile adequate data for calculation of a median.

Regional river and stream TSS medians and professional judgment (considering climatic and topographic characteristics) were used to identify and group regions. Six combined regions were used to develop TSS means for lakes and reservoirs. Combined regions included the East (Mid-Atlantic, South Atlantic-Gulf, Great Lakes, Ohio, and Tennessee), the Mississippi (Upper and Lower Mississippi), and the Midwest (Missouri, Arkansas-White-Red, and Texas-Gulf). Although they had considerably fewer TSS measurements than the combined regions, the Great Basin and California regions were analyzed separately because their waterbodies and climate are too different in character to combine with other regions.

Table 4-9 shows for each region and waterbody type the calculated median values and the years of record and number of measurements used to calculate the median. Note that lakes show significantly lower TSS values than rivers and that patterns in inter-regional variability are similar for the river and lake datasets, supporting the rationale used to aggregate and analyze the STORET TSS data.

4.4 Generating Spatially Averaged Sector-Level Human and Livestock Populations

Human population estimates were used to generate estimates for both cancer and noncancer effects for local populations (i.e., those human populations living within study areas). Livestock population estimates were used to project statistical cancer incidence within the

Table 4-9. Default TSS Values Used in Combustor Risk Analysis

	Hydrologic Region		Years of Record	No. of Measurements	Median TSS (mg/L)
STORET Median TSS - Rivers and Streams					
	2	Mid-Atlantic	60 - 97	47,076	27
	3	South Atlantic-Gulf	60 - 64, 67 - 97	43,013	24
	4	Great Lakes	60 - 96	29,538	21
	5	Ohio	60 - 97	39,899	27
	6	Tennessee	60 - 61, 65, 71, 73 - 96	4,136	15
	7	Upper Mississippi	60 - 96	34,382	68
	8	Lower Mississippi	60 - 97	44,649	163
	10	Missouri	60 - 97	62,767	120
	11	Arkansas-White-Red	60 - 97	46,863	206
	12	Texas-Gulf	60 - 61, 64 - 96	7,268	72
	16	Great Basin	64, 66 - 97	19,930	13
	18	California	60 - 96	41,999	57
STORET Median TSS - Lakes and Reservoirs					
Group					
East	2	Mid-Atlantic	63, 66 - 69, 73 - 93	549	6
	3	South Atlantic-Gulf			
	4	Great Lakes			
	5	Ohio			
	6	Tennessee			
Mississippi	7	Upper Mississippi	60 - 63, 67, 72 - 75, 77, 83, 93, 95 - 96	1,694	38
	8	Lower Mississippi			
Midwest	10	Missouri	60, 62 - 79, 81 - 96	2,142	70
	11	Arkansas-White-Red			
	12	Texas-Gulf			
Great Basin	16	Great Basin	80 - 82, 85	35	1
California	18	California	74 - 79, 88, 90	23	9

general population (i.e., the human population located across the United States) resulting from the ingestion of agricultural commodities that are produced within study areas and impacted by dioxin released from HWC facilities but distributed nationally for consumption. This section describes the data sources and methodologies used to generate the sector-level human and livestock population estimates that were used in the HWC risk analysis.

4.4.1 Human Receptor Populations

Sector-level population projections for human receptors could be generated only for "enumerated receptors" (i.e., those receptor populations for which U.S. Census and Census of Agriculture data could be used to generate sector-level population estimates). The enumerated receptor populations considered in the HWC risk analysis were: residents; home gardeners; and commercial beef, dairy, pork, and produce farmers.

The recreational fishers are an enumerated receptor population but with some important differences from the other receptor populations. The recreational fisher is discussed separately in Section 4.4.1.2.

4.4.1.1 Enumerated Receptor Populations. Estimation of sector-level population totals for enumerated receptor populations involves the use of both 1990 U.S. Census (U.S. EPA, 1995) block-group-level data (U.S. Census data) and 1987/1992 Census of Agriculture (U.S. Department of Commerce, 1993) county-level data (Census of Agriculture data)[7]. The U.S. Census provides detailed population density data, which are broken down into the number of *total persons* and the number of *persons in rural area on farm*. The HWC risk analysis estimates risks for four separate age groups for each receptor population (0-5, 6-11, 12-19, and >19 years). Therefore, U.S. Census data, for both total persons and persons in rural areas on farms, were obtained for each of these four age groups. However, the U.S. Census does not provide a detailed breakdown of the type of agricultural activity for individuals or families (e.g., how many beef cattle farm families or dairy farm families are present in a given census block). Therefore, county-level Census of Agriculture data, which do contain detailed agricultural activity data at the farm level, were used in conjunction with the U.S. census data.

Because individual U.S. Census block groups often do not correspond exactly to the shape of individual sectors within a given study area (e.g., some Census blocks may overlap several sectors while others are contained completely within a given sector), it is often necessary to apportion a given Census block group's population between several sectors. The assumption was made in the HWC risk analysis that the U.S. Census block group populations are evenly distributed within each Census block. Therefore, the proportion of a Census block group that lies within a given sector was used to determine the proportion of that Census block group's population that was apportioned to that sector (see Figure 4-3).

[7] The Census of Agriculture county-level data are collected on a 5-year cycle. The most recent collection efforts (1987 and 1992) did not include 1990, the year when the most recent U.S. Census data were published. Therefore, to match the U.S. Census data with the Census of Agriculture data with regard to year of coverage, Census of Agriculture data from 1987 were averaged together with Census of Agriculture data from 1992 in order to represent 1990.

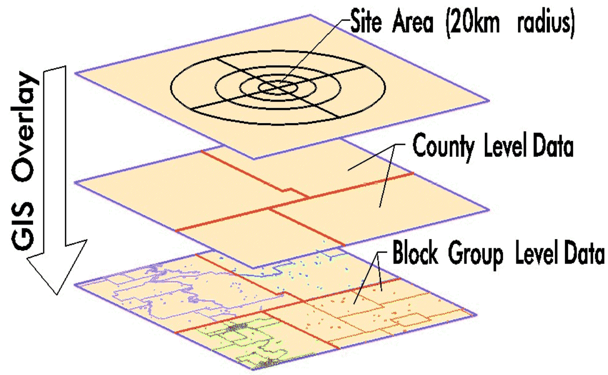

Figure 4-3. Example of inputs to sector/population averaging program.

County-level Census of Agriculture data were then used to further differentiate sector population estimates for total farmers (differentiated into four age groups) into estimates for specific farmer receptor populations (e.g., to convert the number of farmers >19 years of age in a given sector into the number of beef cattle farmers >19 years of age or the number of dairy cattle farmers >19 years of age). Because Census of Agriculture data are available only at the county level and not at the smaller scale block level, the assumption was made that the ratios of *specific farm type* to *total farms* at the county level applied uniformly across the entire county. This assumption allowed the trends in specific farm family ratios (e.g., the percentage of farm families that are dairy farm families) to be applied to all sectors that fall within a given county. When a given sector extended across more than one county, the specific farm family ratios from the different counties were apportioned based on the area proportion of the sector that each county overlapped.

The enumeration of population was conducted for each study area for each facility independent of other HWC facilities regardless of the proximity of these other facilities. Therefore, there are situations in which the 20-km study areas of two or more facilities overlap but the effect of this overlap was not considered. Overall, approximately 15 percent of the individuals residing within HWC study areas are impacted by more than one HWC facilities.

The effect of overlapping study areas on risk results is not known. The aggregate impact of chemical constituents emitted from multiple facilities on individuals residing within overlap areas was not evaluated. Failure to model aggregate impacts for human receptors residing in the area of overlap would underestimate chemical constituent concentrations. On the other hand, the overlap of study areas results in “double counting” of exposed individuals, since the same overlap population is assessed separately for each facility. Because some areas are impacted by more than two facilities, the amount of double counting varies. Those facilities located at industrial facilities in proximity to one another, such as on-site incinerators, overlap with greatest frequency. In aggregate, the percent of double counting is approximately 24 percent.

The following U.S. Census and Census of Agriculture data categories were used to differentiate specific receptor populations for each sector using the methods described above:

- **Residents:** U.S. Census data were used to estimate the number of residents in each of the four age groups of interest. Residents were further differentiated into “residents” and “residents who are home gardeners.” The percentage of residents who are home gardeners was set at 38 percent (U.S. EPA, 1997, *Exposure Factors Handbook* Table 13-1, 1986 Vegetable Gardening by Demographic Factors). This percentage applied to all age groups; that is, children of home gardeners were included in the home gardener population. Nonfarm resident household population estimates were adjusted to exclude nonfarm residents engaged in home gardening so that these two receptor populations were mutually exclusive.

- **Beef cattle farmers:** The total number of individuals on farms, obtained from the U.S. Census data, was adjusted by the ratio of beef farms to total farms obtained from county-level Census of Agriculture data. The specific Census of Agriculture data category used to represent beef farmers was “beef cows (farms)” obtained from Table 1, County Summary Highlights.

- **Dairy cattle farmers:** The total number of individuals on farms, obtained from the U.S. Census data, was adjusted by the ratio of dairy farms to total farms obtained from county-level Census of Agriculture data. The specific Census of Agriculture data category used to represent dairy farmers was “milk cows (farms)” obtained from Table 1, County Summary Highlights.

- **Pork farmers:** The total number of individuals on farms, obtained from the U.S. Census data, was adjusted by the ratio of pork farms to total farms obtained from county-level Census of Agriculture data. The specific Census of Agriculture data category used to represent pork farmers was “hog and pig inventory (farms)” obtained from Table 1, County Summary Highlights.

- **Produce farmers:** The total number of individuals on farms, obtained from the U.S. Census data, was adjusted by the ratio of produce farms to total farms obtained from county-level Census of Agriculture data. The produce receptor is intended to include all individuals engaged in raising exposed fruits/vegetables and root vegetables. Therefore, the following Census of Agriculture data categories were summed to obtain an estimate of the total number of farms raising these crops: “Irish potatoes (farms),” “Veg hv for sale (farms),” “Land in orchards (farms),” and “Dry edible beans, exc dry limas (farms).” Each of these data categories is found in Table 1 of the Census of Agriculture Data, County Summary Highlights.

Individual commercial farmer receptor populations (i.e., beef, dairy, pork, and produce farmers) were determined by multiplying total farm population within a sector by the percentage of total farm population that represents each type of commercial farmer. These percentages were obtained from the U.S. Census of Agriculture, which lists farm types and the percentage of total farmers within the farm type. These U.S. Census of Agriculture percentages total to more than

100 percent of total farmers, indicating that an individual farmer who is engaged in both beef and dairy production is counted in both groups. For the purposes of this risk analysis, counting a farmer in more than one group was consistent with the intent of the commercial farmer receptor populations, which was to determine exposures associated with a particular type of farming activity. In the case of multipurpose farms, the farmer is exposed by more than one route (e.g., beef and dairy) and is counted in both receptor populations. The fact that this farmer was counted separately means that the peak exposure that may result from multipurpose farming activity was not considered. It should also be noted that not all farm categories were considered in this risk analysis and, therefore, commercial farm receptor populations do not total to the number of total farms. Only farm types that represent the most important exposure pathways for the constituents modeled in this risk analysis were considered.

Table 4-10 presents the enumerated population counts for each human receptor population. These receptor population totals are shown for each source category. The population counts in Table 4-10 are facility-weighted values, meaning that they are the total estimated population counts for individuals residing within 20 km of all HWC combustor facilities nationwide.

The sector-level estimates for each receptor population obtained using the methodologies detailed above are combined with sector-specific individual risk estimates for each receptor population to project population risks for a given study area.

4.4.1.2 <u>Recreational Fisher</u>. A key factor in generating sector-level risk estimates for the recreational fisher was the ability to characterize the magnitude of recreational fishing activity at specific modeled waterbodies. Unlike the other enumerated receptor populations, risk for recreational fishers is not primarily dependent on their sector location but rather on which waterbodies they frequent for fishing. Multiple factors influence the level of fishing activity at a specific waterbody including: (1) population density within the study area in which the waterbody is located, (2) accessibility to the waterbody, and (3) the presence of competing waterbodies (with regard to fishing activity) in the vicinity of the waterbody under evaluation. If a relatively larger number of waterbodies favored for recreational fishing activity were located outside of a given study area (but still within a reasonable fishing trip travel distance), then recreational fishers residing within that study area may regularly travel outside of the study area to fish at waterbodies that are less impacted by HWC emissions because those waterbodies are farther from the HWC facility.

Local population risk was characterized for the recreational fisher using semiquantitative risk statements (see Section 6.2.2). These semiquantitative population risk statements required the generation of study-area-level recreational fisher population projections (instead of sector-level projections). This site-characterization task was conducted using 1991 National Survey of Fishing, Hunting, and Wildlife data (U.S. DOI, 1993). Specifically, National Survey data provide county-level values for the fraction of the rural and urban population that engages in recreational fishing activity and is 16 years old or older. These county-level values were applied to each of the U.S. Census block groups within a given study area (block groups are differentiated into urban versus rural categories) to project the number of recreational fishers per block group. The block-group-specific estimates for recreational fishers were then spatially apportioned to the sectors within a given study area in the same manner used for other receptor populations to

Table 4-10. Population Summary by Source Category and Receptor (V2)1

	CK	CINC	LWAK	OINCS	OINCL	Area Sources: CK	Area Sources: INC	All INC
Resident								
0-1 yr	7,942	43,152	6,948	365,994	157,524	261	32,698	566,670
0-5 yr	54,736	291,622	47,753	2,639,476	1,008,948	2,016	213,370	3,940,045
6-11 yr	56,158	290,335	46,083	2,665,349	1,019,878	2,222	223,839	3,975,561
12-19 yr	74,096	372,948	64,333	3,519,595	1,280,145	3,338	292,876	5,172,688
20 yr +	477,459	2,353,478	424,962	23,799,666	8,303,097	16,972	1,702,740	34,456,241
Total	662,450	3,308,383	583,130	32,624,085	11,612,067	24,548	2,432,825	47,544,535
Home Gardner								
0-1 yr	4,868	26,448	4,259	224,319	96,547	160	20,041	347,314
0-5 yr	33,548	178,736	29,268	1,617,743	618,387	1,236	130,775	2,414,866
6-11 yr	34,419	177,947	28,244	1,633,601	625,086	1,362	137,192	2,436,634
12-19 yr	45,414	228,581	39,430	2,157,171	784,605	2,046	179,504	3,170,357
20 yr +	292,636	1,442,454	260,461	14,586,892	5,088,995	10,402	1,043,615	21,118,341
Total	406,018	2,027,718	357,403	19,995,407	7,117,073	15,045	1,491,086	29,140,199
Beef Farmer								
0-1 yr	80	83	23	182	118	9	79	383
0-5 yr	631	661	160	1,256	768	77	621	2,684
6-11 yr	721	806	166	1,481	908	82	786	3,195
12-19 yr	947	920	246	1,857	1,167	119	896	3,945
20 yr +	5,291	5,513	1,466	10,509	6,504	632	5,072	22,526
Total	7,590	7,900	2,038	15,104	9,347	910	7,375	32,350
Dairy Farmer								
0-1 yr	16	15	4	24	28	1	15	67
0-5 yr	122	123	27	171	191	6	122	485
6-11 yr	149	147	28	194	217	6	147	558
12-19 yr	180	171	39	255	268	9	173	694
20 yr +	1,029	1,018	231	1,555	1,511	45	994	4,084
Total	1,480	1,460	325	2,174	2,187	65	1,436	5,821

(continued)

Table 4-10. (continued)

	CK	CINC	LWAK	OINCS	OINCL	Area Sources: CK	Area Sources: INC	All INC
Produce Farmer								
0-1 yr	6	9	2	24	42	0	8	76
0-5 yr	43	80	12	192	311	2	67	584
6-11 yr	51	90	12	198	337	2	77	625
12-19 yr	63	110	17	256	435	3	91	801
20 yr +	363	665	98	1,604	2,471	16	544	4,740
Total	521	946	137	2,249	3,554	23	778	6,749
Pork Farmer								
0-1 yr	33	12	4	78	46	4	11	136
0-5 yr	244	101	25	526	311	31	91	938
6-11 yr	279	121	26	608	362	34	114	1,092
12-19 yr	359	146	38	780	443	49	137	1,370
20 yr +	1,986	861	225	4,536	2,441	270	750	7,838
Total	2,868	1,228	314	6,451	3,558	384	1,092	11,237

generate an estimate of the number of recreational fishers per sector. The estimates for each of the 16 sectors within a given study area were then summed to generate an overall estimate for the total recreational fishers within a given study area. Although these study-area-level population estimates were generated using sector-level calculations, the underlying data do not have sufficient resolution to allow sector-level population inferences to be drawn and used in risk characterization.

4.4.1.3 <u>Enumeration of Human Populations for PM Analysis</u>. To evaluate functions that relate particulate matter to specific health effects for a specific subpopulation (e.g., ages 65 and over), estimates are needed of the number of people in a particular population subgroup who are exposed to a given change in air quality. For this risk assessment, in addition to the receptor populations described in Section 4.4.1.2, it was necessary to develop sector-level population estimates for the variety of population subgroups that were examined by the PM concentration-response functions.

For the PM analysis, sector-level population estimates developed from the U.S. Census were available for two age categories that are commonly examined in the concentration-response functions used in this analysis: ages 18 to 65 and ages 65 and over. For other age groups needed for the PM analysis, the percentage of persons in the subpopulation at the county level was applied to the sector level. For example, Census data are available that estimate that, in Autauga

County, AL, 5 percent of the population is between the ages of 8 and 12 (the population examined by Schwartz et al. (1994) in a study of lower respiratory symptoms). This percentage was then multiplied by the total sector population to estimate the total number of children ages 8 to 12 in a sector that lies completely within Autauga County.

The above-described method is straightforward when a sector lies completely within one county; however, many sectors lie in multiple counties. Sectors lying in more than one county were assigned a spatially weighted average of the county-level subpopulation percentage breakdowns. This spatially weighted average was determined by multiplying the proportion of each sector (in terms of area) located in a given county by that county's subpopulation percentage. The resulting proportion-adjusted county-specific data were then summed for all the counties in which a sector lies, giving an estimate of sector-level subpopulation percentages. This spatially weighted averaging method assumes that county populations are uniformly distributed throughout the county. See Appendix E to obtain a more complete explanation of how the population data were incorporated into the PM analysis.

4.4.2 Livestock Populations

The projection of livestock populations at the sector level also involves integrating U.S. Census block-group-level data with Census of Agriculture county-level data. The U.S. Census data provide detailed estimates of the number of farms located within each sector. These sector-level estimates were modified using adjustment factors derived from county-level Census of Agriculture data to estimate the total number of animals, for livestock animals of interest, located within each sector[8]. The adjustment factors used were

- **Proportion of total farms that are within each specific farm category:** This adjustment factor converts the sector-level total farm numbers into totals for each of the farm categories (e.g., beef farms and dairy farms).

- **Average number of animals located on a single farm:** This adjustment factor allows the number of farms within a given sector to be converted to number of animals for livestock categories of interest within a given sector.

The use of Census of Agriculture data in this manner assumes that trends in the county-level data hold across the entire county and therefore can be applied to all sectors falling within that county.

The assumption was also made that, when a given U.S. Census block group falls within several sectors, the total number of farms within that block group can be apportioned to the different sectors based on the relative portion of that Census block that falls within each sector. This assumption is the same as that used in projecting human receptor populations.

[8] The HWC risk analysis evaluated dioxin cancer risks to the general public from the ingestion of the following food commodities: pork, beef, and milk. Therefore, sector-level livestock population projections were completed for beef cattle, dairy cattle, and hogs.

It is important to note that the Census of Agriculture includes all farms, irrespective of whether there is a house located on the farm, while the U.S. Census data include only those farms containing houses. Because livestock on all farms were of interest, not just farms with houses, the ratio of *total farms*, obtained from the county-level Census of Agriculture data, to *housing units rural (farm)*, obtained from the county-level U.S. Census data, was used to adjust the sector-specific estimates of total farm numbers. This ratio corrects for the fact that the U.S. Census data, which form the basis of the sector-level projections of total farms, do not include farms without houses.

The sector-level estimates for each livestock category obtained were used in assessing national population cancer risk (see Section 8.3.1.2).

4.5 References

Anderson, J.R., E.E. Hardy, J.T. Roach, and R.E. Witmer. 1976. A Land Use and Land Cover Classification System for Use with Remote Sensor Data. Geological Survey Professional Paper 964. In: *U.S. Geological Survey Circular 671*. United States Geological Survey, Washington, DC. http://mapping.usgs.gov/pub/ti/LULC/lulcpp964/lulcpp964.txt.

Keup, L.E. 1985. Flowing water resources. *Water Resources Bulletin* 21(2):291-296.

Laveck, G. and M. Coombs (U.S. Environmental Protection Agency). 1996. Better Assessments Science Integrating Point and Nonpoint Sources (BASINS) CD-ROM. Office of Water, Washington, DC.

Schwartz, J., D.W. Dockery, L.M. Neas, D. Wypij, J.H. Ware, J.D. Spengler, P. Koutrakis, F.E. Speizer, and B.G. Ferris, Jr. 1994. Acute effects of summer air pollution on respiratory symptom reporting in children. *American Journal of Respiratory and Critical Care Medicine* 150:1234-1242. As cited in Abt.

Seaber, P.R., F.P. Kapinos, and G.L. Knapp. 1987. *Hydrologic Unit Maps*. U.S. Geological Survey Water-Supply Paper 2294. U.S. Government Printing Office, Washington, DC. pp. 1 to 13.

U.S. DOC (Department of Commerce). 1993. *1987/1992 Census of Agriculture. Final county file (machine readable file)*. Bureau of the Census, Economics and Statistics Administration, Washington, DC.

U.S. DOI (Department of the Interior). 1993. *1991 National Survey of Fishing, Hunting and Wildlife-Associated Recreation*. U.S. Fish and Wildlife Service, Washington, DC.

U.S. EPA (Environmental Protection Agency). n.d. STORET (STORage and RETrieval system). Office of Water, Washington, DC.

U.S. EPA (Environmental Protection Agency). 1994. U.S. EPA Reach File. Version 3.0 Alpha Release (RF3-Alpha), Technical Reference .

U.S. EPA (Environmental Protection Agency). 1995. 1990 Census Block Group Boundaries in the Conterminous United States. Office of Information Resources Management, Washington, DC.

U.S. EPA (Environmental Protection Agency). 1997. *Exposure Factors Handbook*. EPA/600/P-95/002Fa. Office of Research and Development, Washington, DC.

U.S. EPA (Environmental Protection Agency). 1999a. *Assessment of the Potential Costs, Benefits, and Other Impacts of the Hazardous Waste Combustion MACT Standards*. Washington, DC.

U.S. EPA (Environmental Protection Agency). 1999b. *Final Technical Support Document for HWC MACT Standards, Volume V. Emissions Estimates and Engineering Costs*. Washington, DC.

USDA (U.S. Department of Agriculture). 1971. Chapter 6: Sediment sources, yields, and delivery ratios. In: *National Engineering Handbook, Section 3: Sedimentation,* Soil Conservation Service, Washington, DC. pp. 6-1 to 6-14.

USDA (U.S. Department of Agriculture). 1978. *Predicting Rainfall Erosion Losses: A Guide to Conservation Planning. Agriculture Handbook No. 537*. Science and Education Administration, Washington, DC.

USGS (U.S. Geological Survey). 1994. Water Data Storage and Retrieval System (WATSTORE). National Water Data Exchange, Reston, VA.

Wanielista, M.P., and Y.A. Yousef. 1993. *Stormwater Management*. John Wiley & Sons, Inc., New York, NY. pp. 399 to 410.

Weesies, G.A., Conservation Agronomist, USDA Natural Resources Conservation Service (NRCS), National Soil Erosion Research Laboratory, Purdue University, West Lafayette, IN. 1998. Personal communication with S. Guthrie, Research Triangle Institute (estimation of length and slope factors). June 8.